Remington's Pharmaceutical Sciences

Eighteenth Edition

Remington's Pharmaceutical Sciences

18

18TH
EDITION

Remington's

ALFONSO R GENNARO
Editor, and Chairman
of the Editorial Boord

Pharmaceutical Sciences

1990

MACK PUBLISHING COMPANY
Easton, Pennsylvania 18042

Library of Congress Catalog Card No. 60-53334

ISBN 0-912734-04-3

The use of structural formulas from USAN and the USP Dictionary of Drug Names is by permission of The USP Convention. The Convention is not responsible for any inaccuracy contained herein.

NOTICE—This text is not intended to represent, nor shall it be interpreted to be, the equivalent of or a substitute for the official United States Pharmacopeia (USP) and/or the National Formulary (NF). In the event of any difference or discrepancy between the current official USP or NF standards of strength, quality, purity, packaging and labeling for drugs and representations of them herein, the context and effect of the official compendia shall prevail.

Printed in the United States of America by the Mack Printing Company, Easton, Pennsylvania

Remington's Pharmaceutical Sciences . . . *a treotise on the theory and practice of the pharmaceutical sciences, with essential information about pharmaceutical and medicinal agents; also a guide to the professional responsibilities of the pharmacist as the drug-information specialist of the health team . . . A textbook and reference work for pharmacists, physicians and other practitioners of the pharmaceutical and medical sciences.*

Eighteenth Edition—1990

Published in the 170th year of the
PHILADELPHIA COLLEGE OF PHARMACY AND SCIENCE

Remington Historical / Biographical Data

The following is a record of the editors and the dates of publication of successive editions of this book, prior to the 13th Edition known as *Remington's Practice of Pharmacy* and subsequently as *Remington's Pharmaceutical Sciences.*

First Edition, 1886
Second Edition, 1889
Third Edition, 1897
Fourth Edition, 1905

Joseph P Remington

Fifth Edition, 1907
Sixth Edition, 1917

Joseph P Remington
Assisted by
E Fullerton Cook

Seventh Edition, 1926

Editors
E Fullerton Cook
Charles H LaWall

Eighth Edition, 1936

Editors
E Fullerton Cook
Charles H LaWall

Associate Editors
Ivor Griffith
Adley D Nichols
Arthur Osol

Ninth Edition, 1948

Editors
E Fullerton Cook
Eric W Martin

Tenth Edition, 1951

Editors
E Fullerton Cook
Eric W Martin

Eleventh Edition, 1956

Editors
Eric W Martin
E Fullerton Cook

Associate Editors
E Emerson Leuallen
Arthur Osol
Linwood F Tice
Clarence T Van Meter

Twelfth Edition, 1961

Editors
Eric W Martin
E Fullerton Cook
E Emerson Leuallen
Arthur Osol
Linwood F Tice
Clarence T Van Meter

Assistant to the Editors
John E Hoover

Thirteenth Edition, 1965

Editor-in-Chief
Eric W Martin

Managing Editor
John E Hoover

Editors
Grafton D Chase
Herald R Cox
Richard A Deno
Alfonso R Gennaro
Stewart C Harvey
Robert E King
E Emerson Leuallen
Arthur Osol
Ewart A Swinyard
Clarence T Van Meter

Fourteenth Edition, 1970

Chairmon, Editorial Board
Arthur Osol

Managing Editor
John E Hoover

Editors
Grafton D Chase
Richard A Deno
Alfonso R Gennaro
Melvin R Gibson
Stewart C Harvey
Robert E King
Alfred N Martin
Ewart A Swinyard
Clarence T Van Meter
Bernard Witlin

Fifteenth Edition, 1975

Chairman, Editorial Board
Arthur Osol

Managing Editor
John E Hoover

Editors
John T Anderson
Cecil L Bendush
Grafton D Chase
Alfonso R Gennaro
Melvin R Gibson
C Boyd Granberg
Stewart C Harvey
Robert E King
Alfred N Martin
Ewart A Swinyard

Sixteenth Edition, 1980

Chairman, Editorial Board
Arthur Osol

Editors
Grafton D Chase
Alfonso R Gennaro
Melvin R Gibson
C Boyd Granberg
Stewart C Harvey
Robert E King
Alfred N Martin
Ewart A Swinyard
Gilbert L Zink

Seventeenth Edition, 1985

Chairman, Editorial Board
Alfonso R Gennaro

Editors
Grafton D Chase
Ara Der Marderosian
Stewart Harvey
Daniel A Hussar
Thomas Medwick
Edward G Rippie
Joseph D Schwartz
Ewart A Swinyard
Gilbert L Zink

Editorial Board Members and Editors

Alfonso R Gennaro, PhD / *Philadelphia College of Pharmacy and Science*—Professor of Chemistry. Chairman of the Editorial Board and Editor, *Remington's Pharmaceutical Sciences.* Coauthor, Chapter 22. Coeditor, Part 6, *Pharmaceutical and Medicinal Agents.*

Grofton D Chase, PhD / *Philadelphia College of Pharmacy and Science*—Emeritus Professor of Chemistry, Editor, Part 5, *Radioisotopes in Pharmacy and Medicine.* Author, Chapters 32 and 33.

Aro DerMorderosian, PhD / *Philadelphia College of Pharmacy and Science*—Professor of Pharmacognosy. Research Professor in Medicinal Chemistry. Editor, Part 1, *Orientation.* Author, Chapters 64, 74 and 96.

Stewart C Harvey, PhD / *University of Utah School of Medicine*—Professor of Pharmacology. Editorial Board Member. Editor, Part 6, *Pharmaceutical and Medicinal Agents.* Author, Chapters 35, 38, 40, 43 to 47, 49, 61 and 62. Coauthor, Chapters 36, 41 and 50.

Daniel A Hussar, PhD / *Philadelphia College of Pharmacy and Science*—Remington Professor of Pharmacy. Editorial Board Member. Editor, Part 9, *Pharmaceutical Practice.* Author, Chapters 100 and 102.

Thomas Medwick, PhD / *Rutgers University*—Professor and Chairman, Department of Pharmaceutical Chemistry. Editorial Board member. Editor, Part 3, *Pharmaceutical Chemistry,* and Part 4, *Testing and Analysis.* Coauthor, Chapter 26.

Edward G Rippie, PhD / *University of Minnesota College of Pharmacy*—Professor of Pharmaceutics. Editorial Board Member. Editor, Part 2, *Pharmaceutics.* Author, Chapter 12. Coauthor, Chapter 88.

Joseph D Schwartz, PhD / *Philadelphia College of Pharmacy and Science*—Linwood Tice Professor of Pharmaceutics. Editorial Board Member. Editor, Part 8, *Pharmaceutical Preparations and Their Manufacture.* Coauthor, Chapters 88 and 89.

Ewart A Swinyard, PhD / *University of Utah*—Professor Emeritus of Pharmacology, College of Pharmacy and School of Medicine. Editorial Board member. Editor, Part 6, *Pharmaceutical and Medicinal Agents.* Author, Chapters 39, 42, 48, 53 to 60, 63, 65 and 70. Coauthor, Chapter 66.

Gilbert L Zink, PhD / *Philadelphia College of Pharmacy and Science*—Associate Professor of Biology, Editor, Part 7, *Biological Products.* Author, Chapter 71.

Authors

The following contributors to the Eighteenth Edition of *Remington's Pharmaceutical Sciences* served as authors or coauthors, along with the editors and members of the Editorial Board, of the 109 chapters of this book.

Hamed M Abdou, PhD / Vice President, Worldwide Pharmaceutical Technical Operations, E R Squibb & Sons, Inc; Author of Chapter 30, *Instrumental Methods of Analysis* and Chapter 31, *Dissolution*.

Anu B Amerson, PharmD / Professor, College of Pharmacy/Director, Drug Information Center, Chandler Medical Center, University of Kentucky; Author of Chapter 103, *Clinical Drug Literature*.

Howard C Ansel, PhD / Professor of Pharmacy and Dean, College of Pharmacy, University of Georgia; Author of Chapter 101, *The Prescription*.

Kenneth E Avis, DSc / Emeritus Professor, Pharmaceutics, College of Pharmacy, University of Tennessee, Memphis; Author of Chapter 84, *Parenteral Preparations*.

Leonard C Bailey, PhD / Associate Professor of Pharmaceutical Chemistry, Rutgers University College of Pharmacy; Author of Chapter 29, *Chromatography*.

Lawrence H Block, PhD / Professor of Pharmaceutics, Duquesne University School of Pharmacy; Author of Chapter 87, *Medicated Applications*.

Joseph B Bogardus, PhD / Basic Pharmaceutics Research, Bristol-Myers Company; Coauthor of Chapter 18, *Reaction Kinetics*.

Sanford Bolton, PhD / Chairman, Department of Pharmacy and Administrative Sciences, St John's University; Author of Chapter 10, *Statistics*.

John Bosso, PharmD / Professor of Clinical Pharmacy and Adjunct Professor of Pediatrics, College of Pharmacy and School of Medicine, University of Utah; Coauthor of Chapter 34, *Diseases: Manifestations and Pathophysiology*.

B Sue Brizuela, MS / Assistant Professor of Information Science, Head of Public Services, Joseph W England Library, Philadelphia College of Pharmacy and Science; Coauthor of Chapter 7, *Drug Information*.

Dale B Christensen, PhD / Associate Professor, Department of Pharmacy Practice, School of Pharmacy, University of Washington; Coauthor of Chapter 11, *Computer Science*.

Sebastian G Ciancio, DDS / Professor and Chairman, Department of Periodontology, School of Dental Medicine, State University of New York at Buffalo; Author of Chapter 109, *Dental Services*.

Kenneth A Connors, PhD / Professor of Pharmaceutics, School of Pharmacy, University of Wisconsin; Author of Chapter 14, *Complex Formation*.

Anthony J Cutie, PhD / Professor of Pharmaceutics, Arnold and Marie Schwartz College of Pharmacy and Health Sciences, Long Island University; Coauthor of Chapter 92, *Aerosols*.

Anthony R DiSanto, PhD / Vice President, Drug Delivery Research and Development, The Upjohn Company; Author of Chapter 76, *Bioavailability and Bioequivalency Testing*.

Clarence A Discher,* PhD / Professor Emeritus, Rutgers University; Author of Chapter 21, *Inorganic Pharmaceutical Chemistry*.

Clyde R Erskine, Jr, BSc / Vice President, Corporate Quality Audits and Services, SmithKline Beckman Corporation; Author of Chapter 82, *Quality Assurance and Control*.

Lorraine D Evans, BS, H(ASCP) / Clinical Pathology, Bristol-Myers Company; Coauthor of Chapter 28, *Clinical Analysis*.

William E Fassett, BS, MBA / Assistant Professor, Department of Pharmacy Practice, School of Pharmacy, University of Washington; Coauthor of Chapter 11, *Computer Science*.

Joseph L Fink III, BS(Pharm), JD / Assistant Dean and Professor, College of Pharmacy, University of Kentucky; Coauthor of Chapter 107, *Laws Governing Pharmacy*.

Michael R Franklin, PhD / Professor of Pharmacology, College of Pharmacy and School of Medicine, University of Utah; Author of Chapter 52, *Enzymes*.

Ruta Freimanis, BS, RPh / Associate Secretary, United States Adopted Names Council; Coauthor of Chapter 24, *Drug Nomenclature—United States Adopted Names*.

James W Freston, MD, PhD / Professor and Chairman, Department of Medicine, University of Connecticut Health Center; Coauthor of Chapter 34, *Diseases: Manifestations and Pathophysiology*.

Robert L Giles, BA / Vice President and General Manager, Glenn Beall Engineering Inc; Coauthor of Chapter 80, *Plastic Packaging Materials*.

Harold N Godwin, MS / Professor and Director of Pharmacy, The University of Kansas Medical Center; Author of Chapter 94, *Institutional Patient Care*.

Frederick J Goldstein, PhD / Professor of Pharmacology, Philadelphia College of Pharmacy and Science; Coauthor of Chapter 69, *Pharmacological Aspects of Substance Abuse*.

A Richard Goolkasian, BS, RPh / Director of Alumni and Professional Affairs, Massachusetts College of Pharmacy and Allied Health Sciences; Author of Chapter 1, *Scope*.

Gerald Hecht, PhD / Director Process Development, Alcon Laboratories; Coauthor of Chapter 86, *Ophthalmic Preparations*.

Judith A Hesp, MS / Instructor in Information Science, Coordinator of Bibliographic Instruction, Joseph W England Library, Philadelphia College of Pharmacy and Science; Coauthor of Chapter 7, *Drug Information*.

Gregory J Higby, PhD / Director, American Institute of the History of Pharmacy, School of Pharmacy, University of Wisconsin-Madison; Author of Chapter 2, *Evolution of Pharmacy*.

Andrew S Katocks, Jr, PhD / Senior Research Pharmacologist, American Cyanamid Company, Medical Research Division; Coauthor of Chapter 27, *Biological Testing*.

Calvin H Knowlton, MDiv, RPh / Clinical Associate Professor of Pharmacy, Philadelphia College of Pharmacy and Science; Author of Chapter 4, *The Practice of Community Pharmacy*.

Richard W Knueppel, RPh / President, Knueppel Home Health Care Center; Author of Chapter 104, *Health Accessories*.

Harry B Kostenbauder, PhD / Associate Dean for Research, College of Pharmacy, University of Kentucky; Coauthor of Chapter 18, *Reaction Kinetics*.

Richard L Kronenthal, PhD / Director of Research, Ethicon Inc; Author of Chapter 105, *Surgical Supplies*.

Arthur J Lawrence, PhD, RPh / Office of the Assistant Secretary of Health, US Public Health Service; Author of Chapter 6, *Pharmacists in Government*.

Eric J Lien, PhD / Professor of Pharmacy / Pharmaceutics and Biomedical Chemistry, School of Pharmacy, University of Southern California; Author of Chapter 13, *Molecular Structure, Properties and States of Matter*.

Mark A Longer, PhD / MCR Research Fellow, Department of Biological Sciences, University of Keele; Coauthor of Chapter 91, *Sustained-Release Drug Delivery Systems*.

Werner Lowenthal, PhD / Professor of Pharmacy and Pharmaceutics and Professor of Educational Development and Planning, School of Pharmacy, Medical College of Virginia; Author of Chapter 9, *Metrology and Calculation*.

Karen B Main, PhD / Physical Pharmacist, Pharmaceutical Development Department, ICI Pharmaceuticals Group; Coauthor of Chapter 26, *Analysis of Medicinals*.

Duane D Miller, PhD / Professor and Chairman, Division of Medicinal Chemistry and Pharmacognosy, College of Pharmacy,

* Deceased

The Ohio State University; Author of Chapter 25, *Structure-Activity Relationship and Drug Design.*

Michael Montagne, PhD / Associate Professor of Pharmacy Administration, Philadelphia College of Pharmacy and Science; Coauthor of Chapter 3, *Ethics* and Author of Chapter 99, *Drug Education.*

John D Mullins, PhD / Consultant; Coauthor of Chapter 86, *Ophthalmic Preparations.*

Maven J Myers, PhD / Professor of Pharmacy Administration, Philadelphia College of Pharmacy and Science; Coauthor of Chapter 3, *Ethics.*

J G Nairn, PhD / Professor of Pharmacy, Faculty of Pharmacy, University of Toronto; Author of Chapter 83, *Solutions, Emulsions, Suspensions and Extracts.*

Paul J Niebergall, PhD / Professor of Pharmaceutical and Sciences / Director, Pharmaceutical Development Center, Medical University of South Carolina; Author of Chapter 17, *Ionic Solutions and Electrolytic Equilibria.*

Robert E O'Connor, PhD / Merck Frosst Canada, Inc; Coauthor of Chapter 88, *Pawders.*

Melanie O'Neill / Becton Dickinson & Company; Coauthor of Chapter 78, *Sterilization.*

Richard W Pecina, PhD / President, Richard W Pecina & Associates; Coauthor of Chapter 80, *Plastic Packaging Materials.*

Garnet E Peck, PhD / Professor of Industrial Pharmacy / Director of the Industrial Pharmacy Laboratory, Purdue University; Author of Chapter 77, *Separation.*

G Briggs Phillips, PhD / Becton Dickinson & Company; Coauthor of Chapter 78, *Sterilization.*

Nicholas G Popovich, PhD / Associate Professor of Pharmacy Practice, School of Pharmacy and Pharmacal Sciences, Purdue University; Author of Chapter 93, *Ambulatory Patient Care.*

Stuart C Porter, PhD / Vice President, Research and Development, Colorcon; Author of Chapter 90, *Coating of Pharmaceutical Dosage Forms.*

Galen Radebaugh, PhD / Director of Pharmaceutics, Parke-Davis Pharmaceutical Research Division, Warner-Lambert Company; Coauthor of Chapter 75, *Preformulation.*

Paul L Ranelli, PhD / Assistant Professor of Pharmacy Administration, Philadelphia College of Pharmacy and Science; Author of Chapter 98, *Patient Communication.*

Louis J Ravin, PhD / Department of Pharmaceutics, Research and Development, Smith Kline & French Laboratories; Coauthor of Chapter 75, *Preformulation.*

Jack W Reich, PhD / Vice President Regulatory Affairs, Gensia Pharmaceuticals, Inc; Author of Chapter 8, *Research.*

James W Richards, MBA / Professor of Pharmacy Administration, College of Pharmacy, University of Michigan; Author of Chapter 108, *Community Pharmacy Economics and Management.*

Jack Robbins, PhD / Director, Pharmacy Affairs, Schering Laboratories; Author of Chapter 5, *Opportunities for Pharmacists in the Pharmaceutical Industry.*

Joseph R Robinson, PhD / Professor of Pharmacy, School of Pharmacy, University of Wisconsin; Coauthor of Chapter 91, *Sustained-Release Drug Delivery Systems.*

Frank Roia, PhD / Professor of Biology, Philadelphia College of Pharmacy and Science; Author of Chapter 72, *Immunizing Agents and Diagnostic Skin Antigens.*

Douglas E Rollins, MD, PhD / Associate Professor of Medicine and Pharmacology, School of Medicine and College of Pharmacy, University of Utah; Author of Chapter 37, *Clinical Pharmacokinetics.*

G Victor Rossi, PhD / Vice Presidont of Academic Affairs / Professor of Pharmacology, Philadelphia College of Pharmacy and Science; Coauthor of Chapter 27, *Biological Testing* and Coauthor of Chapter 69,

Edward Rudnic, PhD / Director, Formulation Development, Schering Research; Coauthor of Chapter 89, *Oral Solid Dosage Forms.*

Donald O Schiffman, PhD / Secretary, United States Adopted Names Council; Coauthor of Chapter 24, *Drug Nomenclature—United States Adopted Names.*

Hans Schott, PhD / Professor of Pharmaceutics and Colloid Chemistry, School of Pharmacy, Temple University; Coauthor of Chapter 19, *Disperse Systems* and Author of Chapter 20, *Rheology.*

John J Sciarra, PhD / President, Retail Drug Institute / Professor of Industrial Pharmacy, Arnold and Marie Schwartz College of Pharmacy and Health Sciences, Long Island University; Coauthor of Chapter 92, *Aerosols.*

John H Shinkai, PhD / Emeritus Professor of Pharmaceutical Chemistry, Rutgers University, College of Pharmacy; Coauthor of Chapter 22, *Organic Pharmaceutical Chemistry.*

E Richard Shough, PhD / Associate Dean and Professor, University of Oklahoma Health Sciences Center, College of Pharmacy; Author of Chapter 73, *Allergenic Extracts.*

Frederick P Siegel, PhD / Professor of Pharmaceutics, College of Pharmacy, University of Illinois; Author of Chapter 79, *Tonicity, Osmoticity, Osmolality and Osmolarity.*

Larry M Simonsmeier, BS(Pharm), JD / Associate Dean and Professor, College of Pharmacy, Washington State University; Coauthor of Chapter 107, *Laws Governing Pharmacy.*

Robert D Smyth, PhD / Vice President, Pharmaceutical Development, Bristol-Myers Company; Coauthor of Chapter 28, *Clinical Analysis.*

Thomas C Snader, PharmD / Consultant Pharmacist; Author of Chapter 95, *Long-Term Care Facilities.*

Theodore D Sokoloski, PhD / Professor of Pharmacy, College of Pharmacy, The Ohio State University; Author of Chapter 16, *Solutions and Phase Equilibria.*

Robert B Stewart, MS / Professor and Chairman, Department of Pharmacy Practice, College of Pharmacy, University of Florida; Author of Chapter 67, *Adverse Drug Reactions.*

James Swarbrick, DSc, PhD / Professor and Chairman, Division of Pharmaceutics, School of Pharmacy, University of North Carolina at Chapel Hill; Coauthor of Chapter 19, *Disperse Systems.*

Anthony R Temple, MD / Director, Regulatory and Medical Affairs, McNeil Consumer Products Company / Adjunct Associate Professor, Department of Pediatrics, University of Pennsylvania School of Medicine / Lecturer, Philadelphia College of Pharmacy and Science; Author of Chapter 106, *Paison Control.*

John P Tischio, PhD / Principle Scientist, Immunohiology Research Institute; Author of Chapter 68, *Pharmacogenetics.*

Salvatore J Turco, PharmD / Professor of Pharmacy, Temple University School of Pharmacy; Author of Chapter 85, *Intravenous Admixtures.*

Elizabeth B Vadas, PhD / Merck Frosst Canada, Inc; Author of Chapter 81, *Stability of Pharmaceutical Products.*

Ernestine Vanderveen, PhD / National Institute on Drug Abuse, ADAMHA; Coauthor of Chapter 51, *Vitamins and Other Nutrients.*

John E Vanderveen, PhD / Division of Nutrition, Food and Drug Administration; Coauthor of Chapter 51, *Vitamins and Other Nutrients.*

Vincent S Venturella, PhD / Section Manager, Pharmaceutical Research, Anaquest, Div of BOC; Author of Chapter 23, *Natural Products.*

Albert I Wertheimer, PhD / Professor and Director, Department of Graduate Studies in Social and Administrative Pharmacy, College of Pharmacy, University of Minnesota; Author of Chapter 97, *The Patient: Behavioral Determinants.*

Timothy S Wiedmann, PhD / Assistant Professor, College of Pharmacy, University of Minnesota; Author of Chapter 15, *Thermodynamics.*

C Dean Withrow, PhD / Associate Professor of Pharmacology, School of Medicine, University of Utah; Coauthor of Chapter 36, *Basic Pharmacokinetics,* Coauthor of Chapter 41, *Cardiovascular Drugs* and Coauthor of Chapter 50, *Hormones.*

George Zografi, PhD / Professor, School of Pharmacy, University of Wisconsin; Coauthor of Chapter 19, *Disperse Systems.*

Preface to the First Edition

The rapid and substantial progress made in Pharmacy within the last decade has created a necessity for a work treating of the improved apparatus, the revised processes, and the recently introduced preparations of the age.

The vast advances made in theoretical and applied chemistry and physics have much to do with the development of pharmaceutical science, and these have been reflected in all the revised editions of the Pharmacopoeias which have been recently published. When the author was elected in 1874 to the chair of Theory and Practice of Pharmacy in the Philadelphia College of Pharmacy, the outlines of study which had been so carefully prepared for the classes by his eminent predecessors, Professor William Procter, Jr, and Professor Edward Parrish, were found to be not strictly in accord, either in their arrangement of the subjects or in their method of treatment. Desiring to preserve the distinctive characteristics of each, an effort was at once made to frame a system which should embody their valuable features, embrace new subjects, and still retain that harmony of plan and proper sequence which are absolutely essential to the success of any system.

The strictly alphabetical classification of subjects which is now universally adopted by pharmacopoeias and dispensatories, although admirable in works of reference, presents an effectual stumbling block to the acquisition of pharmaceutical knowledge through systematic study; the vast accumulation of facts collected under each head being arranged lexically, they necessarily have no connection with one another, and thus the saving of labor effected by considering similar groups together, and the value of the association of kindred subjects, are lost to the student. In the method of grouping the subjects which is herein adopted, the constant aim has been to arrange the latter in such a manner that the reader shall be gradually led from the consideration of elementary subjects to those which involve more advanced knowledge, whilst the groups themselves are so placed as to follow one another in a natural sequence.

The work is divided into six parts. Part I is devoted to detailed descriptions of apparatus and definitions and comments on general pharmaceutical processes.

The Official Preparations alone are considered in Part II. Due weight and prominence are thus given to the Pharmacopoeia, the National authority, which is now so thoroughly recognized.

In order to suit the convenience of pharmacists who prefer to *weigh solids* and *measure liquids*, the official formulas are expressed, in addition to parts by weight, in *avoirdupois weight* and *apothecaries' measure*. These equivalents are printed in *bold type* near the margin, and arranged so as to fit them for quick and accurate reference.

Part III treats of Inorganic Chemical Substances. Precedence is of course given to official preparation in these. The descriptions, solubilities, and tests for identity and impurities of each substance are systematically tabulated under its proper title. It is confidently believed that by this method of arrangement the valuable descriptive features of the Pharmacopoeia will be more prominently developed, ready reference facilitated, and close study of the details rendered easy. Each chemical operation is accompanied by equations, whilst the reaction is, in addition, explained in words.

The Carbon Compounds, or Organic Chemical Substances, are considered in Part IV. These are naturally grouped according to the physical and medical properties of their principal constituents, beginning with simple bodies like cellulin, gum, etc, and progressing to the most highly organized alkaloids, etc.

Part V is devoted to Extemporaneous Pharmacy. Care has been taken to treat of the practice which would be best adapted for the needs of the many pharmacists who conduct operations upon a moderate scale, rather than for those of the few who manage very large establishments. In this, as well as in other parts of the work, operations are illustrated which are conducted by manufacturing pharmacists.

Part VI contains a formulary of Pharmaceutical Preparations which have not been recognized by the Pharmacopoeia. The recipes selected are chiefly those which have been heretofore rather difficult of access to most pharmacists, yet such as are likely to be in request. Many private formulas are embraced in the collection; and such of the preparations of the old Pharmacopoeias as have not been included in the new edition, but are still in use, have been inserted.

In conclusion, the author ventures to express the hope that the work will prove an efficient help to the pharmaceutical student as well as to the pharmacist and the physician. Although the labor has been mainly performed amidst the harassing cares of active professional duties, and perfection is known to be unattainable, no pains have been spared to discover and correct errors and omissions in the text. The author's warmest acknowledgments, are tendered to Mr A B Taylor, Mr Joseph McCreery, and Mr George M Smith for their valuable assistance in revising the proof sheets, and to the latter especially for his work on the index. The outline illustrations, by Mr John Collins, were drawn either from the actual objects or from photographs taken by the author.

Philadelphia, October, 1885 JPR.

Preface to the Eighteenth Edition

In anticipation of setting forth this *Preface* and prior to gathering thoughts on paper (or more accurately, the word processor!), this Editor paused to reread the preface to the first edition of *Remington*, published in 1885. Since it appears on the preceding page of this book it is recommended highly. The first paragraph would he just as suitable today as penned by Professor Remington 105 years ago.

Each decade transcends the previous one and the pharmaceutical and health sciences are not laggards. Every revision of *Remington* has encompassed new viewpoints, ideas, doctrines or principles which, perhaps, were inconceivable for the previous edition. It is a tribute to the authors and editors that they have kept abreast of the burgeoning literature in their respective fields of expertise.

Change not withstanding, the organization of this edition is similar to its immediate predecessors, being divided into 9 Parts, each subdivided into several chapters. Every chapter has been culled, revised and rewritten to update the material presented.

Two new chapters are evident; *Biotechnology and Drugs* (Chapter 74) and *Drug Education* (Chapter 99). Three chapters of the previous edition, which embraced *Interfacial and Particle Phenomena* and *Colloidal and Coarse Dispersions* have been winnowed and combined into a single chapter entitled, *Disperse Systems* (Chapter 19).

The current revision contains an additional 21 pages. A large amount of space (about 19 pages) gleaned from the review and condensation process, coupled with the extra pages, have been devoted primarily to expanding the contents of Part 6, *Pharmaceutical and Medicinal Agents* and Part 9, *Pharmaceutical Practice*.

Excessive duplication of text is the bane of any editor dealing with a multitude of authors. While some duplication in the discussion of rudimentary concepts is beneficial, there has been a special effort to cross-reference and eliminate unnecessary repetition. Space is at such a premium that it is hoped the reader will not be offended by being diverted to a different section of the text in order to obtain supplementary information.

Photographs which depicted the typical "black box" have been eliminated almost completely and replaced by line drawings or schematic diagrams which are instructive rather than picturesque.

Most of the drug monographs have been revamped not only as a means of updating, but to gain a degree of uniformity. All structural formulas are now in the standard *USAN* form. Duplication of chemical names has been minimized and the inclusion of trade names increased. No attempt has been made to ferret out every trade name by which a product is known, and only the most common are mentioned. The standard format for the major monographs is: Official Name, chemical name (*CAS*—inverted), trade name(s) and manufacturer(s), structural formula, *CAS* (*Chemical Abstracts System*) registry number (in brackets), molecular formula and formula weight (in parenthesis). This is followed by the method of preparation (or a reference if the method is lengthy), physical description, solubility, uses, dose and dosage forms.

The number of authors remains at 97, however, 36 new authors have joined as contributors to *Remington*. As the credentials of the new authors touch upon many areas of pharmacy, every section of the book has been invigorated by the incorporation of updated and fresh concepts.

As one primarily responsible for the production of a comprehensive text devoted to the science and practice of pharmacy, the wisdom of Dr Eric Martin, editor of the 13th Edition, in creating an Editorial Board to share the enormous burden, has been evident constantly. Each of the section editors labored diligently to comply with the logistics of maintaining a smooth flow of manuscripts and proofs. Also, each section editor doubled as an author or coauthor of one or more chapters. It would be remiss not to extend special mention to this group of dedicated people.

Four members of the Editorial Board are serving for the first time after having been authors for several editions. Dr Ara DerMarderosian of PCP&S, Editor for Part 1; Dr Daniel Hussar, also of PCP&S, Part 9; Dr Edward Rippie of the University of Minnesota, Part 2; and Dr Joseph Schwartz of PCP&S, Part 8. Each of the new members literally "jumped into the fray," gave much of their precious time and have become "blooded" members of the staff.

The stalwarts of the Editorial Board surviving the tribulations of one or more previous editions of this work demand singular attention. Dr Grafton Chase of PCP&S for Part 5, *Radioisotopes in Pharmacy & Medicine*; Dr Thomas Medwick of Rutgers University for Part 3, *Pharmaceutical Chemistry* and Part 4, *Testing and Analysis*; and Dr Gilbert Zink of PCP&S for Part 7, *Biological Products*.

Two dauntless, prolific contributors claim special recognition. Drs Stewart Harvey and Ewart Swinyard, both of the University of Utah, have served on the Editorial Board for twenty and twenty-five years respectively. They bear the burden of Part 6, *Pharmaceutical & Medicinal Agents*, which comprises over one-third of the book. Their diligence and meticulous attention to detail has eased the task of this Editor. Our relationship over the past several decades has been one of exceptional pleasure.

The Mack Publishing Company, through Messers Paul Mack and David Palmer, continues its unrelenting support, which has endured through many, many editions of this publication. Special commendation must be extended to Ms Nancy Smolock, of the Mack organization, as she was the person who interfaced with the Editorial Board. She was competent, cooperative and much too tolerant of the many requests made of her.

As with any publication a few of the editorial staff bear the hrunt of the unglamorous, but absolutely essential, chores associated with the production of this voluminous tome. It mandates a close working relationship and, at times, restraint and concession to sustain the harmony necessary to function efficiently. One often encounters the aphorism usually attributed to administrators, "When three managers meet to discuss a problem there arise four points-of-view." Fortunately, this dilemma did not surface in the association of this Editor with Mr John Hoover and Ms Bonnie Packer.

After shepherding this publication through four editions, the Twelfth to Fifteenth, following a short hiatus for the Sixteenth, Mr Hoover returned in a lesser capacity with the Seventeenth revision. With the current edition he reassumes the role of Managing Editor and his experience in pharmacy, journalism and the publishing business, have provided the capabilities needed to translate a disarranged manuscript into a format acceptable hy the publisher and pleasing to the reader.

Ms Packer accepted the assignment of scrutinizing every word of text in the proof stages. Combining her skills in the

health and social sciences, she assumed the charge of reading primarily for comprehension and clarity of presentation, while concurrently uncovering typographical, spelling and grammatical errors which, although unpardonable, are ever-present. As a consequence of her deliberations, passages were often rephrased and refined to portray a concept from the viewpoint of the student, for whom this work primarily is directed.

The *Index* was developed by Mr Hoover, with the assistance of Ms Packer. Much use was made of the computer in ensuring that a complete, practical and useful index was created. It is the opinion of this Editor that a major weakness encountered in most reference books is a perfunctory, casual index which amounts to little more than an expanded table of contents. Users of the index of this book will find it "friendly."

Philadelphia, February, 1990 ARG

Table of Contents

Index

CHAPTER 16

Solutions and Phase Equilibria

Theodore D Sokoloski, PhD
Professor of Pharmacy, College of Pharmacy
Ohio State University
Columbus, OH 43210

Solutions and Solubility

A solution is a chemically and physically homogeneous mixture of two or more substances. The term solution generally denotes a homogeneous mixture that is liquid even though it is possible to have homogeneous mixtures which are solid or gaseous. Thus, it is possible to have solutions of solids in liquids, liquids in liquids, gases in liquids, gases in gases and solids in solids. The first three of these are most important in pharmacy and ensuing discussions will be concerned primarily with them.

In pharmacy different kinds of liquid dosage forms are used and all consist of the dispersion of some substance or substances in a liquid phase. Depending on the size of the dispersed particle they are classified as *true solutions*, *colloidal solutions* or *suspensions*. If sugar is dissolved in water, it is supposed that the ultimate sugar particle is of molecular dimensions and that a *true solution* is formed. On the other hand, if very fine sand is mixed with water, a *suspension* of comparatively large particles, each consisting of many molecules, is obtained. Between these two extremes lie *colloidal solutions*, the dispersed particles of which are larger than those of true solutions but smaller than the particles present in suspensions. In this chapter only true solutions will be discussed.

It is possible to classify broadly all solutions as one of two types.

In the first type, although there may be a lesser or greater interaction between the dispersed substance (the solute) and the dispersing medium (the solvent), the solution phase contains the same chemical entity as found in the solid phase and, thus, upon removal of the solvent, the solute is recovered unchanged. One example would be sugar dissolved in water where, in the presence of sugar in excess of its solubility, there is an equilibrium between sugar molecules in the solid phase with sugar molecules in the solution phase. A second example would be dissolving silver chloride in water. Admittedly, the solubility of this salt in water is low, but it is finite. In this case the solvent contains silver and chloride ions and the solid phase contains the same material. The removal of the solvent yields initial solute.

In the second type the solvent contains a compound which is different from that in the solid phase. The difference between the compound in the solid phase and solution is due generally to some chemical reaction that has occurred in the solvent. An example would be dissolving aspirin in an aqueous solvent containing some basic material capable of reacting with the acid aspirin. Now the species in solution would not only be undissociated aspirin, but aspirin also as its anion, whereas the species in the solid phase is aspirin in only its undissociated acid form. In this situation, if the solvent were removed, part of the substance obtained (the salt of aspirin) would be different from what was present initially in the solid.

Solutions of Solids in Liquids

Reversible Solubility without Chemical Reaction—From a pharmaceutical standpoint solutions of solids in liquids, with or without accompanying chemical reaction in the solvent, are of the greatest importance, and many quantitative data on the behavior and properties of such solutions are available. This discussion will be concerned with definitions of solubility, the rate at which substances go into solution and with temperature and other factors which control solubility.

Solubility—When an excess of a solid is brought into contact with a liquid, molecules of the former are removed from its surface until equilibrium is established between the molecules leaving the solid and those returning to it. The resulting solution is said to be saturated at the temperature of the experiment, and the extent to which the solute dissolves is referred to as its *solubility*. The extent of solubility of different substances varies from almost imperceptible amounts to relatively large quantities, but for any given solute the solubility has a constant value at constant temperature.

Under certain conditions it is possible to prepare a solution containing a larger amount of solute than is necessary to form a saturated solution. This may occur when a solution is saturated at one temperature, the excess of solid solute removed and the solution cooled. The solute present in solution, even though it may be less-soluble at the lower temperature, does not always separate from the solution and there is produced a *supersaturated solution*. Such solutions, formed by sodium thiosulfate or potassium acetate, for example, may be made to deposit their excess of solute by vigorous shaking, scratching the side of the vessel in contact with the solution or introducing into the solution a small crystal of the solute.

Methods of Expressing Solubility—When quantitative data are available, solubilities may be expressed in many ways. For example, the solubility of sodium chloride in water at 25° may be stated as

1. 1 g of sodium chloride dissolves in 2.786 mL of water. (An approximation of this method is used by the USP.)
2. 35.89 g of sodium chloride dissolves in 100 mL of water.
3. 100 mL of a saturated solution of sodium chloride in water contains 31.71 g of solute.
4. 100 g of a saturated solution of sodium chloride in water contains 26.47 g of solute.
5. 1 L of a saturated solution of sodium chloride in water contains 5.425 moles of solute. This also may be stated as a saturated solution of sodium chloride in water is 5.425 molar with respect to the solute.

In order to calculate *3* from *1* or *2* it is necessary to know the density of the solution, in this case 1.198 g/mL. To calculate *5*, the number of grams of solute in 1000 mL of solution (obtained by multiplying the data in (*3*) by ten) is divided by the molecular weight of sodium chloride, namely, 58.45.

Several other concentration expressions are used. Molality is the number of moles of solute in 1000 g of solvent and could be calculated from the data in *4* by subtracting grams

The author acknowledges the kind assistance of Dr Gordon L Flynn, University of Michigan, in the revision of parts of this chapter.

Table I—Descriptive Terms for Solubility

Descriptive Terms	Parts of Solvent for 1 Part of Solute
Very soluble	Less than 1
Freely soluble	From 1 to 10
Soluble	From 10 to 30
Sparingly soluble	From 30 to 100
Slightly soluble	From 100 to 1000
Very slightly soluble	From 1000 to 10,000
Practically insoluble, or insoluble	More than 10,000

of solute from grams of solution to obtain grams of solvent, relating this to 1000 g of solvent and dividing by molecular weight to obtain moles.

Mole fraction is the fraction of the total number of moles present which are moles of one component. Mole % may be obtained by multiplying mole fraction by 100. Normality refers to the number of gram equivalent weights of solute dissolved in 1000 mL of solution.

In pharmacy, use also is made of three other concentration expressions. Percent by weight (% *w/w*) is the number of grams of solute per 100 g of solution and is exemplified by *4* above. Percent weight in volume (% *w/v*) is the number of grams of solute per 100 mL of solution and is exemplified by *3* above. Percent by volume (% *v/v*) is the number of milliliters of solute in 100 mL of solution, referring to solutions of liquids in liquids. The USP indicates that the term "percent," when unqualified, means percent weight in volume for solutions of solids in liquids and percent by volume for solutions of liquids in liquids.

When, in pharmacopeial texts, it has not been possible, or in some instances not desirable, to indicate exact solubility, a descriptive term has been used. Table I indicates the meaning of such terms.

Rate of Solution—It is possible to define quantitatively the rate at which a solute goes into solution. The simplest treatment is based on a model depicted in Fig 16-1. A solid particle dispersed in a solvent is surrounded by a thin layer of solvent having a finite thickness, l in cm. The layer is an integral part of the solid and, thus, is referred to characteristically as the "stagnant layer." This means that regardless of how fast the bulk solution is stirred the stagnant layer remains a part of the surface of the solid, moving wherever the particle moves. The thickness of this layer may get smaller as the stirring of the bulk solution increases, but it is important to recognize that this layer will always have a finite thickness however small it may get.

Using Fick's First Law of Diffusion the rate of solution of the solid can be explained, in the simplest case, as the rate at which a dissolved solute particle diffuses through the stagnant layer to the bulk solution. The driving force behind the movement of the solute molecule through the stagnant layer is the difference in concentration that exists between the concentration of the solute, C_1, in the stagnant layer at the surface of the solid and its concentration, C_2, on the farthest side of the stagnant layer (see *Diffusion in Liquids*, page 221). The greater this difference in concentration ($C_1 - C_2$), the faster the rate of solution.

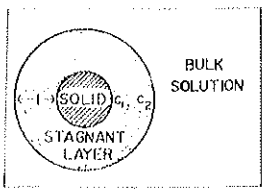


Fig 16-1. Physical model representing the dissolution process.

According to Fick's Law, the rate of solution also is directly proportional to the area of the solid, A in cm^2, exposed to solvent and inversely proportional to the length of the path through which the dissolved solute molecule must diffuse. Mathematically, then, the rate of solution of the solid is given by

$$\text{Rate of solution} = \frac{DA}{l}(C_1 - C_2) \qquad (1)$$

where D is a proportionality constant called the diffusion coefficient in cm^2/sec. In measuring the rate of solution experimentally, the concentration C_2 is maintained at a low value compared to C_1 and hence considered to have a negligible effect on the rate. Furthermore, C_1 most often is the saturation solubility of the solute. Hence Eq 1 is simplified to

$$\text{Rate of solution} = \frac{DA}{l}(\text{saturation solubility}) \qquad (2)$$

Equation 2 quantitatively explains many of the phenomena commonly observed that affect the rate at which materials dissolve.

1. Small particles go into solution faster than large particles. For a given mass of solute, as the particle size becomes smaller, the surface area per unit mass of solid increases; Eq 2 shows that as area increases, the rate must increase proportionately. Hence, if a pharmacist wishes to increase the rate of solution of a drug, its particle size should be decreased.
2. Stirring a solution increases the rate at which a solid dissolves. This is because the thickness of the stagnant layer depends on how fast the bulk solution is stirred; as stirring rate increases, the length of the diffusional path decreases. Since the rate of solution is proportional inversely to the length of the diffusional path, the faster the solution is stirred, the faster the solute will go into solution.
3. The more soluble the solute, the faster is its rate of solution. Again, Eq 2 predicts that the larger the saturation solubility, the faster the rate.
4. With a viscous liquid the rate of solution is decreased. This is because the diffusion coefficient is proportional inversely to the viscosity of the medium; the more viscous the solvent, the slower the rate of solution.

Heat of Solution and Temperature Dependency—Turning from the kinetic aspects of dissolution, this discussion will be concerned with the situation where there is thermodynamic equilibrium between solute in its solid phase and the solute in solution. (It is assumed that there is an amount of solid material in excess of the amount that can go into solution; hence, a solid phase is always present.) As defined earlier, the concentration of solute in solution at equilibrium is the saturation solubility of the substance.

When a solid (Solute A) dissolves in some solvent two steps may be considered as occurring: the solid absorbs energy to become a liquid and then the liquid dissolves.

$$A_{(\text{solid})} \rightleftarrows A_{(\text{liquid})} \rightleftarrows A_{(\text{solution})}$$

For the overall dissolution the equilibrium existing between solute molecules in the solid and solute molecules in solution may be treated as any equilibrium. Thus, for Solute A in equilibrium with its solution

$$A_{(\text{solid})} \rightleftarrows A_{(\text{solution})}$$

Using the Law of Mass Action an equilibrium constant for this system can be defined, just as any equilibrium constant may be written as

$$K_{eq} = \frac{a_{(\text{solution})}}{a_{(\text{solid})}}$$

where a denotes the activity of the solute in each phase. Since the activity of a solid is defined as unity

$$K_{eq} = a_{(\text{solution})}$$

Because the activity of a compound in dilute solution is approximated by its concentration and, because this concentration is the saturation solubility, K_S, the van't Hoff Equation (for a more complete treatment, see Ref 1, page 113) may be used, which defines the relationship between an equilibrium constant (here, solubility) and absolute temperature.

$$\frac{d \log K_S}{dT} = \frac{\Delta H}{2.3RT^2} \tag{3}$$

where $d \log K_S/dT$ is the change of $\log K_S$ with a unit change of absolute temperature, T; ΔH is a constant which in this situation is the heat of solution for the overall process (solid $\rightleftarrows$ liquid $\rightleftarrows$ solution); and R is the gas constant, 1.99 cal/mole/deg. Equation 3, a differential, may be solved to give

$$\log K_S = -\frac{\Delta H}{2.3RT} + J \tag{4}$$

where J is a constant. A more useful form of this equation is

$$\log \frac{K_{S,T_2}}{K_{S,T_1}} = \frac{\Delta H(T_2 - T_1)}{2.3RT_1T_2} \tag{5}$$

where K_{S,T_1} is the saturation solubility at absolute temperature T_1 and K_{S,T_2} is the solubility at temperature T_2. Through the use of Eq 5, if ΔH and the solubility at one temperature are known, the solubility at any other temperature can be calculated.

Effect of Temperature—As is evident from Eq 4, the solubility of a solid in a liquid depends on the temperature. If, in the process of solution, heat is absorbed (as evidenced by a reduction in temperature), ΔH is by convention positive and the solubility of the solute will increase with increasing temperature. Such is the case for most salts, as is shown in Fig 16-2, in which the solubility of the solute is plotted as the ordinate and the temperature as the abscissa, and the line joining the experimental points represents the solubility curve for that solute.

If a solute gives off heat during the process of solution (as evidenced by an increase in temperature), ΔH is, by convention, negative and solubility decreases with an increase in temperature. This is the case with calcium hydroxide and, at higher temperatures, with calcium sulfate. (Because of the slight solubility of these substances their solubility curves are not included.) When heat is neither absorbed nor given off, the solubility is not affected by variation of temperature as is nearly the case with sodium chloride.

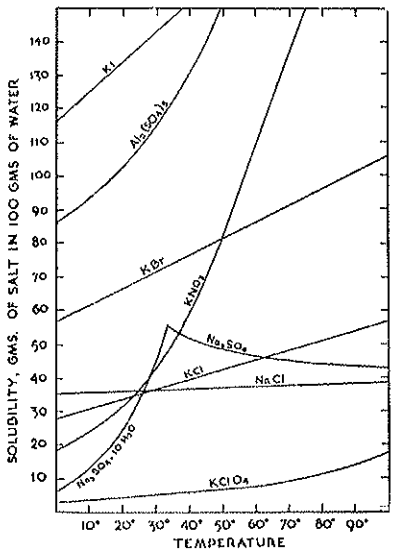


Fig 16-2. Effect of heat on solubility.

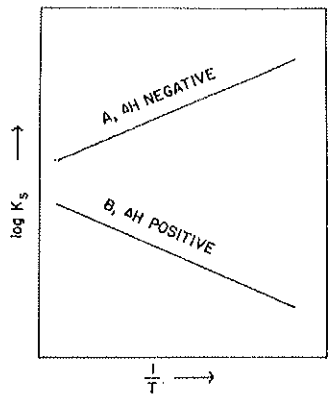


Fig 16-3. Typified relationship between the logarithm of the saturation solubility and the reciprocal of the absolute temperature.

Solubility curves usually are continuous as long as the chemical composition of the solid phase in contact with the solution remains unchanged, but if there is a transition of the solid phase from one form to another, a break will be found in the curve. Such is the case with $Na_2SO_4 \cdot 10H_2O$, which dissolves with absorption of heat up to a temperature of 32.4°, at which point there is a transition of the solid phase to anhydrous sodium sulfate, Na_2SO_4, which dissolves with evolution of heat. This change is evidenced by increased solubility of the hydrated salt up to 32.4°, but above this temperature the solubility decreases.

These temperature effects are what would be predicted from Eq 4. When the heat of solution is negative, signifying that energy is released during dissolution, the relation between $\log K_S$ and $1/T$ is typified in Fig 16-3 (Curve A), where as $1/T$ increases, $\log K_S$ increases. It can be seen that with increasing temperature (T itself actually increases proceeding left in Fig 16-3, A) there is a decrease in solubility. On the other hand, when the heat of solution is positive—that is, when heat is absorbed in the solution process—the relation between $\log K_S$ and $1/T$ is typified in Fig 16-3, B. Here, as temperature increases ($1/T$ decreases), the solubility increases.

Effect of Salts—The solubility of a nonelectrolyte, in water, either is decreased or increased generally by the addition of an electrolyte; it is only rarely that the solubility is not altered. When the solubility of the nonelectrolyte is decreased, the effect is referred to as *salting-out*; if it is increased, it is described as *salting-in*. Inorganic electrolytes commonly decrease solubility, though there are some exceptions to the generalization.

Salting-out occurs because the ions of the added electrolyte interact with water molecules and, thus, in a sense reduce the amount of water available for solution of the nonelectrolyte. (Refer to the section on *Thermodynamics of the Solution Process*, page 215, for another view.) The greater the degree of hydration of the ions, the more the solubility of the nonelectrolyte is decreased. If, for example, one compares the effect of equivalent amounts of lithium chloride, sodium chloride, potassium chloride, rubidium chloride and cesium chloride (all of which belong to the family of alkali metals and are of the same valence type), it is observed that lithium chloride decreases the solubility of a nonelectrolyte to the greatest extent and that the salting-out effect decreases in the order given. This is also the order of the degree of hydration of the cations; lithium ion, being the smallest ion and, therefore, having the greatest density of positive charge per unit of surface area (see also Chapter 13 under *Electronegativity Values*), is the most extensively

hydrated of the cations while cesium ion is hydrated the least. Salting-out is encountered frequently in pharmaceutical operations.

Salting-in, commonly occurs when either the salts of various organic acids or organic-substituted ammonium salts are added to aqueous solutions of nonelectrolytes. In the first case the solubilizing effect is associated with the anion and in the second, with the cation. In both cases the solubility increases as the concentration of added salt is increased. The solubility increase may be relatively great, sometimes amounting to several times the solubility of the nonelectrolyte in water.

Solubility of Solutes Containing Two or More Species—In cases where the solute phase consists of two or more species (as in an ionizable inorganic salt), when the solute goes into solution, the solution phase often contains each of these species as discrete entities. For some such substance, *AB*, the following relationship for the solution process may be written.

$$AB_{(solid)} \rightleftharpoons A_{(solution)} + B_{(solution)}$$

Since there is an equilibrium between the solute and saturated solution phases, the Law of Mass Action defines an equilibrium constant, K_{eq}

$$K_{eq} = \frac{a_{A(solution)} \cdot a_{B(solution)}}{a_{AB(solid)}} \quad (6)$$

where $a_{A(solution)}$, $a_{B(solution)}$ and $a_{AB(solid)}$ are the activities of *A* and *B* in solution and of *AB* in the solid phase. Recall from the earlier discussion that the activity of a solid is defined as unity, and that in a very dilute solution (eg, for a slightly soluble salt), concentrations may be substituted for activities and Eq 6 then becomes

$$K_{eq} = C_A C_B$$

where C_A and C_B are the concentrations of *A* and *B* in solution. In this situation K_{eq} has a special name, the *solubility product*, K_{SP}. Thus

$$K_{SP} = C_A C_B \quad (7)$$

This equation will hold true theoretically only for slightly soluble salts.

As an example of this type of solution, consider the solubility of silver chloride

$$K_{SP} = [Ag^+][Cl^-]$$

where the brackets [ ] designate concentrations.

At 25° the solubility product has a value of 1.56×10^{-10}, the concentration of silver and chloride ions being expressed in moles/liter. The same numerical value applies also to solutions of silver chloride containing an excess of either silver or chloride ions. If the silver-ion concentration is increased by the addition of a soluble silver salt, the chloride-ion concentration must decrease until the product of the two concentrations again is equal numerically to the solubility product. In order to effect the decrease in chloride-ion concentration, silver chloride is precipitated and, hence, its solubility is decreased. In a similar manner an increase in chloride-ion concentration by the addition of a soluble chloride effects a decrease in the silver-ion concentration until the numerical value of the solubility product is attained. Again, this decrease in silver-ion concentration is brought about by the precipitation of silver chloride.

The solubility of silver chloride in a saturated aqueous solution of the salt may be calculated by assuming that the concentration of silver ion is the same as the concentration of chloride ion, both expressed in moles/liter, and that the concentration of dissolved silver chloride is numerically the same since each silver chloride molecule gives rise to one silver ion and one chloride ion. Since

$$[\text{dissolved AgCl}] = [Ag^+] = [Cl^-]$$

the solubility of AgCl is equal to $\sqrt{1.56 \times 10^{-10}}$, which is 1.25×10^{-5} mole/liter. Multiplying this by the molecular weight of silver chloride (143) we obtain a solubility of approximately 1.8 mg/liter.

For a salt of the type $PbCl_2$ the solubility product expression takes the form

$$[Pb^{2+}][Cl^-]^2 = K_{SP}$$

while for As_2S_3 it would be

$$[As^{3+}]^2[S^{2-}]^3 = K_{SP}$$

because from the Law of Mass Action

$$PbCl_{2(solid)} \rightleftharpoons Pb^{2+}{}_{(solution)} + 2Cl^-{}_{(solution)}$$

and

$$As_2S_{3(solid)} \rightleftharpoons 2As^{3+}{}_{(solution)} + 3S^{2-}{}_{(solution)}$$

For further details of methods of using solubility-product calculations, the reader is referred to books on qualitative or quantitative analysis or physical chemistry.

Recall that the solubility-product principle is valid for aqueous solutions of slightly soluble salts, provided the concentration of added salt is not too great. Where the concentrations are high, deviations from the theory occur and these have been explained by assuming that in such solutions the nature of the solvent has been changed. Frequently, deviations also may occur as the result of the formation of complexes between the two salts. An example of increased solubility, by virtue of complex-ion formation, is seen in the effect of solutions of soluble iodides on mercuric iodide. According to the solubility-product principle it might be expected that soluble iodides would decrease the solubility of mercuric iodide, but because of the formation of the more soluble complex salt K_2HgI_4 which dissociates as follows

$$K_2HgI_4 \rightleftharpoons 2K^+ + (HgI_4)^{2-}$$

the iodide ion no longer functions as a common ion.

Practical applications of the solubility-product principle are found in qualitative and quantitative analysis whenever an excess of a precipitant is added in order to diminish, by common-ion effect, the solubility of the precipitate.

It is possible to formulate some general rules regarding the effect of the addition of soluble salts to slightly soluble salts where the added salt does not have an ion common to the slightly soluble salt. If the ions of the added soluble salt are not highly hydrated (see *Effect of Salts on the Solubility of Nonelectrolytes*, page 209), the solubility product of the slightly soluble salt will increase because the ions of the added salt tend to decrease the interionic attraction between the ions of the slightly soluble salt. On the other hand, if the ions of the added soluble salt are hydrated, water molecules become less available and the interionic attraction between the ions of the slightly soluble salt increases with a resultant decrease in solubility product. Another way of considering this effect is discussed later (*Thermodynamics of the Solution Process*, page 215).

The effect of temperature is, in general, what would be expected; increasing the temperature of the solution results in an increase of the solubility product.

Solubility Following a Chemical Reaction—Thus far in this chapter the discussion has been concerned with solubility that comes about because of interplay of entirely physical forces. The dissolution of some substance resulted from overcoming the physical interactions between solute mole-

cules and solvent molecules by the energy produced when a solute molecule interacted physically with a solvent molecule. The solution process, however, can be facilitated also by a chemical reaction. Almost always the chemical enhancement of solubility in aqueous systems is due to the formation of a salt following an acid–base reaction.

An alkaloidal base, or any other nitrogenous base of relatively high molecular weight, generally is slightly soluble in water, but if the pH of the medium is reduced by addition of acid, the solubility of the base is increased, considerably so, as the pH continues to be reduced. The reason for this increase in solubility is that the base is converted to a salt, which is relatively soluble in water. Conversely, the solubility of a salt of an alkaloid or other nitrogenous base is reduced as pH is increased by addition of alkali.

The solubility of slightly soluble acid substances is, on the other hand, increased as the pH is increased by addition of alkali, the reason again being that a salt, relatively soluble in water, is formed. Examples of acid substances whose solubility is thus increased are aspirin, theophylline and the penicillins, cephalosporins and barbiturates. Conversely, the soluhility of salts of the same substances is decreased as the pH decreases.

Among some inorganic compounds a somewhat similar behavior is observed. Tribasic calcium phosphate, $Ca_3(PO_4)_2$, for example, is almost insoluble in water, but if an acid is added its solubility increases rapidly with a decrease in pH. Tbis is because hydrogen ions have such a strong affinity for phosphate ions forming nonionized phosphoric acid, that the calcium phosphate is dissolved in order to release phosphate ions. Or, stated in another way, the solubilization is an example of a reaction in which a strong acid (the source of the hydrogen ions) displaces a weak acid.

In all of these examples soluhilization occurs as tbe result of an interaction of the solute with an acid or a base and that the species in solution is *not* the same as the undissolved solute. Compounds which do not react with either acids or bases are slightly, or not at all, influenced in their aqueous solubility by variations of pH. Such effects as may he observed are generally due to ionic *salt effects*.

It is possible to analyze quantitatively the solubility following an acid–base reaction by considering it as a two-step process. The first example is an organic acid, designated as *HA*, that is relatively insoluble in water. Its two-step dissolution can be represented as

$$HA_{(solid)} \rightleftarrows HA_{(solution)}$$

followed by

$$HA_{(solution)} \rightleftarrows H^+_{(solution)} + A^-_{(solution)}$$

The equilibrium constant for the first step is the solubility of *HA* ($K_S = [HA]_{solution}$), just as was developed earlier when no chemical reaction took place, and the equilibrium constant for the second step is the dissociation constant of the acid is

$$K_a = \frac{[H^+][A^-]}{[HA]}$$

Since the total amount of compound *in solution* is the sum of nonionized and ionized forms of the acid, the total solubility may be designated as $S_{t(HA)}$, or

$$S_{t(HA)} = [HA] + [A^-] = [HA] + K_a \frac{[HA]}{[H^+]} \quad (8)$$

Since $K_S = [HA]$, Eq 8 becomes

$$S_{t(HA)} = K_S\left(1 + \frac{K_a}{[H^+]}\right) \quad (9)$$

Equation 9 is very useful since it equates the total solubility of an acid drug with the hydrogen-ion concentration of the solvent. If the water solubility, K_S, and the dissociation constant, K_a, are known, the total solubility of the acid can be calculated at various hydrogen-ion concentrations. Equation 9 demonstrates quantitatively how the total solubility of the acid increases as the hydrogen-ion concentration decreases (that is, as the pH increases).

It is possible to develop an equation similar to Eq 9 for the solubility of a basic drug, *B*, such as a relatively insoluble nitrogenous base (an alkaloid, for example), at various hydrogen-ion concentrations. The solubility of the base in water may be represented in two steps, as

$$B_{(solid)} \rightleftarrows B_{(solution)}$$
$$B_{(solution)} \rightleftarrows BH^+_{(solution)} + OH^-_{(solution)}$$

Again, if K_S is the solubility of the free hase in water and K_b is its dissociation constant

$$K_b = \frac{[BH^+][OH^-]}{[B]}$$

the total solubility of the hase in water $S_{t(B)}$ is given by

$$S_{t(B)} = [B] + [BH^+] = [B] + \frac{K_b[B]}{[OH^-]} = K_S\left(1 + \frac{K_b}{[OH^-]}\right) \quad (10)$$

It is convenient to rewrite Eq 10 in terms of hydrogen-ion concentration by making use of the dissociation constant for water

$$K_W = [H^+][OH^-] = 1 \times 10^{-14}$$

Equation 10 tben becomes

$$S_{t(B)} = K_S\left(1 + \frac{K_b}{K_W/[H^+]}\right) = K_S\left(1 + \frac{K_b[H^+]}{K_W}\right) \quad (11)$$

Equation 11 quantitatively shows how the total solubility of the base increases as the hydrogen-ion concentration of the solvent increases. If K_S and K_b are known, it is possihle to calculate the total solubility of a basic drug at various hydrogen-ion concentrations using this equation.

Equations 9 and 11 have assumed that the salt formed following a chemical reaction is infinitely soluble. This, of course, is not an acceptable assumption as suggested and demonstrated by Kramer and Flynn.[2] Rather, for an acidic or basic drug there should be a pH at which *maximum solubility* occurs where this soluhility remains the sum of the solution concentrations of the free and salt forms of the drug at that pH. Using a basic drug, B, as the example, this would mean that a solution of B, at pH values greater than the pH of maximum solubility, would be saturated with free-base form but not with the salt form and the use of Eq 11 would be valid for the prediction of solubility. On the other hand, at pH values less than the pH of maximum soluhility, the solution would be saturated with salt form and Eq 11 is no longer really valid. Since in this situation the total solubility of the base, $S_{t(B)}$, is

$$S_{t(B)} = [B] + [BH^+]_s$$

where the subscript, *s*, designates a solution saturated with salt, the correct equation to use at pH values less than the pH maximum would be

$$S_{t(B)} = [BH^+]_s\left(1 + \frac{[OH^-]}{K_b}\right) = [BH^+]_s\left(1 + \frac{K_W}{K_b[H^+]}\right) \quad (12)$$

A relationship similar to Eq 12 likewise can be developed for an acidic drug at pHs greater than its pH of maximum solubility.

Effecting Solution of Solids in the Prescription Laboratory—The method usually employed by the pharmacist when soluble compounds are to be dissolved in water in compounding a prescription requires the use of the mortar and pestle. The ordinary practice is to crush the substance into fragments in the mortar with the pestle and pour the solvent on it, meanwhile stirring with the pestle until solution is effected. If definite quantities are used, and the whole of the solvent is required to dissolve the given weight of the salt, only a portion of the solvent should be added first and, when this is saturated, the solution is poured off and a fresh portion of solvent added. This operation is repeated until the solid is dissolved entirely and all the portions combined. Other methods of effecting solution are to shake the solid with the liquid in a bottle or flask or to apply heat to the substances in a suitable vessel. Substances vary greatly in the rate at which they dissolve; some are capable of producing a saturated solution quickly, others require several hours to attain saturation. All too often, in their haste to prepare a saturated solution, pharmacists fail to obtain the required degree of solution of solute.

With hygroscopic substances like pepsin, silver protein compounds and some others, the best method of effecting solution in water is to place the substance directly upon the surface of the water and then stir vigorously with a glass rod. If the ordinary procedure, such as using a mortar and pestle, is employed with these substances, gummy lumps are formed which are exceedingly difficult to dissolve.

The *solubility* of chemicals and the *miscibility* of liquids are important physical factors for the pharmacist to know, as they often have a bearing on intelligently and properly filling prescriptions. Mainly for the information of the pharmacist, the USP provides tabular data indicating the degree of solubility or miscibility of many official substances.

Determination of Solubility—For the pharmacist and pharmaceutical chemist the question of solubility is of paramount importance. Not only is it necessary to know solubilities when preparing and dispensing medicines, but such information is necessary to effect separation of substances in qualitative and quantitative analysis. Furthermore, the accurate determination of the solubility of a substance is one of the best methods for determining its purity.

The details of the determination of the solubility are affected markedly by the physical and chemical characteristics of the solute and solvent and also by the temperature at which the solubility is to be determined. Accordingly, it is not possible to describe a universally applicable method but, in general, the following rules must be observed in solubility determinations.

1. The purity of both the dissolved substance and the solvent is essential, since impurities in either affect the solubility.
2. A constancy of temperature must be maintained accurately during the course of the determination.
3. Complete saturation must be attained.
4. Accurate analysis of the saturated solution and correct expression of the results are imperative.

Consideration should be given also to the varying rates of solution of different compounds and to the marked effect of the degree of fineness of the particles on the time required for the saturation of the solution.

Many of the solubility data of USP have been determined with regard to the exacting requirements mentioned above.

Phase-Solubility Analysis—This procedure is one of the most useful and accurate methods for the determination of the purity of a substance. It involves the application of precise solubility methods to the principle that constancy of solubility, in the same manner as constancy of melting point, indicates that a material is pure or free from foreign admixture. It is important to recognize that the technique can be used to obtain the exact solubility of the pure substance without the necessity of the experimental material itself being pure.

The method is based on the thermodynamic principles of heterogeneous equilibria which are among the soundest of theoretical concepts of chemistry. Thus, it does not depend on any assumptions regarding kinetics or structure of matter, but is applicable to all species of molecules, and is sufficiently sensitive to distinguish between optical isomers. The requirements for an analysis are simple, since the equipment needed is basic to most laboratories and the quantities of substances required are small.

The standard solubility method consists of five steps:

1. Mixing, in separate systems, increasing amounts of a substance with measured amounts of a solvent.
2. Establishment of equilibrium for each system at identical constant temperature and pressure.
3. Separation of the solid phase from the solutions.
4. Determination of the concentration of the material dissolved in the various solutions.
5. Plotting the concentration of the dissolved material per unit of solvent (*y*-axis, or solution concentration) against the mass of material per unit of solvent (*x*-axis or system concentration).

The solubility method has been established on the sound theoretical principles of the Gibbs phase rule: $F = C - P + 2$, which relates C, the number of components, F, the degrees of freedom (pressure, temperature and concentration) and P, the number of phases for a heterogeneous equilibrium. Since solubility analyses are carried out at constant temperature and pressure, a pure solid in solution would show only one degree of freedom, because only one phase is present at concentrations below saturation. This is represented by section *AB* in Fig 16-4. For a pure solid in a saturated solution at equilibrium (Fig 16-4, *BC*), two phases are present, solid and solution; there is no variation in concentration and thus, at constant temperature and pressure, no degrees of freedom.

The curve *ABC* of Fig 16-4 represents the type of solubility diagram obtained for: (1) a pure material, (2) equal amounts of two or more materials having identical solubilities or (3) a mixture of two or more materials present in the

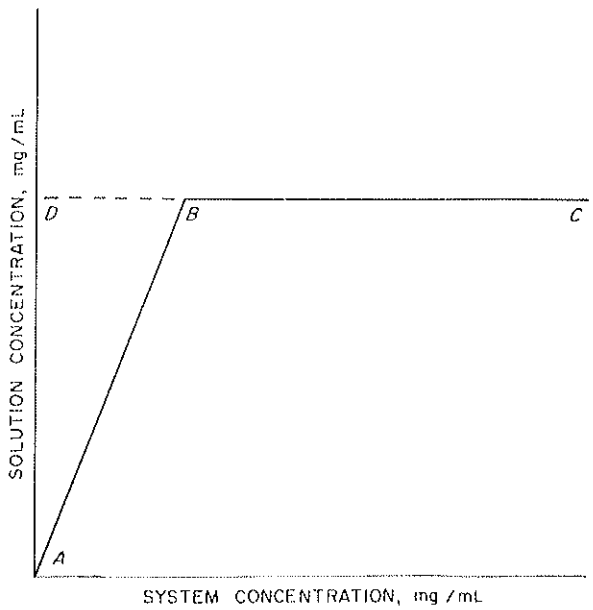


Fig 16-4. Phase-solubility diagram for a pure substance.

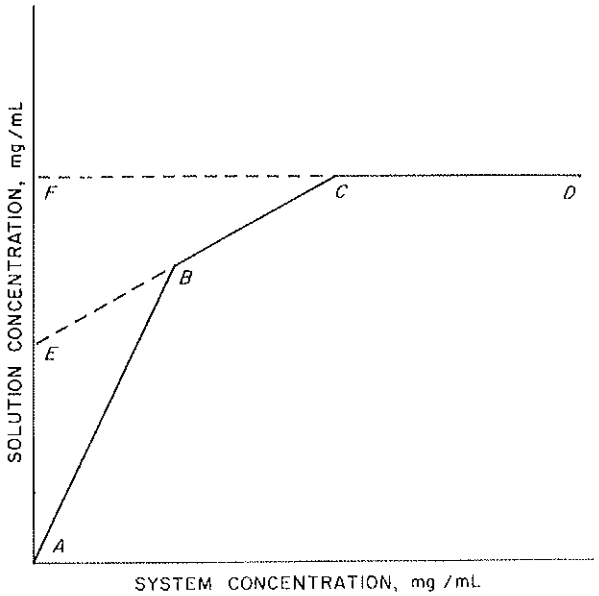


Fig 16-5. Type of solubility curve obtained when a substance contains one impurity.

unique ratio of their solubilities. These latter two cases are rare and often may be detected by a change in solvent system.

Line segment *BC* of Fig 16-4, since it has no slope, usually indicates purity. If, however, this section *does* exhibit a slope, its numerical value indicates the fraction of impurity present. Line segment *BC*, extrapolated to the *y* axis at *D*, is the actual solubility of the pure substance.

A representative type of solubility curve, which is obtained when a substance contains one impurity, is illustrated in Fig 16-5. Here, at *B*, the solution becomes saturated with one component. From *B* to *C* there are two phases present: a solution saturated with Component I (usually the major component) containing also some Component II (usually the minor component), and a solid phase of Component I. The one degree of freedom revealed by the slope of the line segment *BC* is the concentration of Component II, which is the impurity (usually the minor component). A mixture of *d* and *l* isomers could have such a curve, as would any simple mixtures in which the solubilities are independent of each other.

The section *CD* indicates that the solvent is saturated with both components of the two-component mixture. Here, three phases are present: a solution saturated with both components and the two solid phases. No variation of concentration is possible, hence, no degree of freedom is possible (indicated by the lack of slope of section *CD*). The distance *AE* on the ordinate represents the solubility of the major component, and the distance *EF*, the solubility of the minor component.

The fact that the equilibration process is time-consuming, requiring as long as 3 weeks in certain cases, is offset by the fact that all of the sample can be recovered after a determination. This adds to the general usefulness of the method, particularly in cases where the substance is expensive or difficult to obtain. A use for the method other than the determination of purity or of solubility is to obtain especially pure samples by recovering the solid residues at system concentration corresponding to points on section *BC* in Fig 16-5. Thus, the method is useful not only as a quantitative analytical tool, but also for purification.

Solutions of Liquids in Liquids

Binary Systems—Under this title the following types of liquid-pairs may be recognized.

1. Those which are soluble completely in each other in all proportions. Examples: alcohol and water; glycerin and water; alcohol and glycerin.
2. Those which are soluble in each other in definite proportions. Examples: phenol and water; ether and water; nicotine and water.
3. Those which are imperceptibly soluble in each other in any proportion. Examples: castor oil and water; liquid petrolatum and water.

The mutual solubility of liquid pairs of Type 2 has been studied extensively and found to show interesting regularities. If a series of tubes containing varying, but known, percentages of phenol and water are heated (or cooled, if necessary) just to the point of formation of a homogeneous solution, and the temperatures at such points noted, there will be obtained, upon plotting the results, a curve similar to that in Fig 16-6.[3] On this graph the area inside the curve represents the region where mixtures of phenol and water will separate into two layers, while in the region outside of the curve homogeneous solutions will be obtained. The maximum temperature on this curve is called the *critical solution temperature*, that is, the temperature above which a homogeneous solution occurs regardless of the composition of the mixture. For phenol and water the critical solution temperature occurs at a composition of 34.5% phenol in water.

Temperature versus composition curves, as depicted in Fig 16-6, provide much useful information in the preparation of homogeneous mixtures of substances showing mutual-solubility behavior. At room temperature (here assumed to be 25°), by drawing a line parallel to the abscissa at 25°, we find that we actually can prepare two sets of homogeneous solutions, one containing from 0 to about 7.5% phenol and the other containing phenol from 72 to about 95% (its limit of solubility). At compositions between 7.5 and 72% phenol at 25° two liquid layers or phases will separate. In sample tubes containing a concentration of phenol in this two-layer region at 25° one layer always will be phenol-rich and always contain 72% phenol while the other layer will be water-rich and always contain 7.5% phenol. These values are obtained by interpolation of the two points of intersection of the line drawn at 25° with the experimental curve. As it may be deduced, at other temperatures, the composition of the two layers in the two-layer region is determined by the points of intersection of the curve with a line (called

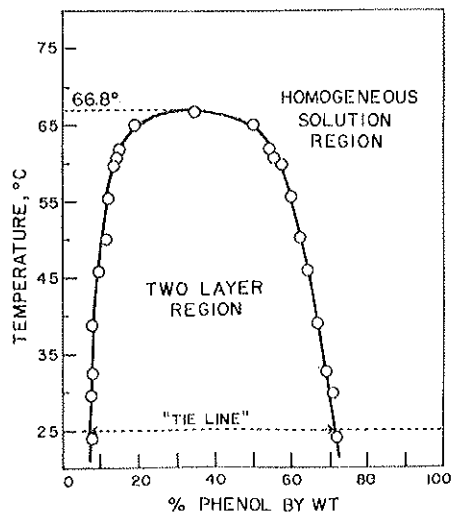


Fig 16-6. Phenol-water solubility.[3]

the "tie line") drawn parallel to the abscissa at that temperature. The relative amounts of the two layers or phases, phenol-rich and water-rich in this example, will depend on the concentration of phenol added. As expected, the proportion of phenol-rich layer relative to the water-rich layer increases as the concentration of phenol added increases. For example, at 20% phenol in water at 25° there would be more of the water-rich layer than of the phenol-rich layer, while at 50% phenol in water there would be more of the phenol-rich layer. The relative proportion of each layer may be calculated from such tie lines at any temperature and composition as well as the amount of phenol present in each of the two phases. To determine how these calculations are made and for further discussion of this topic the student should consult Ref 1, page 79.

A simple and practical advantage in the use of phase diagrams is pointed out in Ref 1. Based on diagrams such as Fig 16-6, they point out that the most concentrated stock solution of phenol that perhaps should be used by pharmacists is one containing 76% *w/w* phenol in water (equivalent to 80% *w/v*). At room temperature this mixture is a homogeneous solution and will remain homogeneous to around 3.5°, at which temperature freezing occurs. It should be noted that Liquefied Phenol USP contains 90% *w/w* phenol and freezes at 17°C. This means that if the storage area in the pharmacy falls to about 63°F, the preparation will freeze, resulting in a stock solution no longer convenient to use.

In the case of phenol and water the mutual solubility increases with increase in temperature and the critical solution temperature occurs at a relatively high point. In a certain number of cases, however, the mutual solubility increases with decrease in temperature and the critical solution temperature occurs at a relatively low value. Most of the substances which show lower critical solution temperatures are amines as, for example, triethylamine with water.

In addition to pairs of liquids which show *either* upper or lower critical solution temperatures, there are other pairs which show *both* upper and lower critical solution temperatures and the mutual solubility curve is of the closed type. An example of this type of liquid pair is found in the case of nicotine and water (Fig 16-7). Mixtures of nicotine and water represented by points within the curve will separate into two layers, but mixtures represented by points outside of the curve are perfectly miscible with each other.

In a discussion of solutions of liquids in liquids it is evident that the distinction between the terms solute and solvent loses its significance. For example, in a solution of water and glycerin, which shall be considered to be the solute and which the solvent? Again, when two liquids are soluble only partially in each other the distinction between solute and solvent might be reversed easily. In such cases the term solvent usually is given to the constituent present in larger quantity.

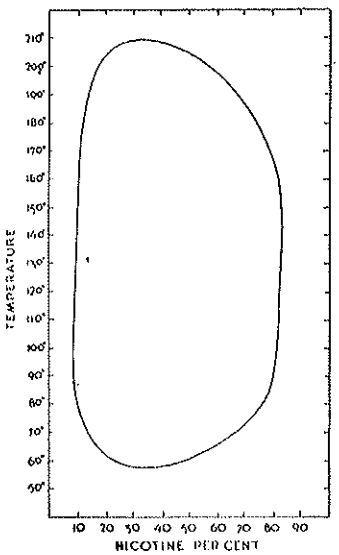


Fig 16-7. Nicotine-water solubility.

Ternary Systems—The addition of a third liquid to a binary liquid system to produce a ternary or three-component system can result in several possible combinations.

If the third liquid is soluble in only one of the two original liquids or if its solubility in the two original liquids is markedly different, the mutual solubility of the original pair will be decreased. An upper critical solution temperature will be elevated and a lower critical solution temperature lowered. On the other hand, the addition of a liquid having roughly the same solubility in both components of the original pair will result in an increase in their mutual solubility. An upper critical solution temperature then will be lowered and a lower critical solution temperature elevated.

An equilateral-triangle graph may be used to represent the situation in which a third liquid is added to a partially miscible liquid pair, the third liquid being miscible with each member of the original pair. In this type of graph, each side of the triangle represents 0% of one of the components and the apex opposite that side represents 100% of that component. The reader is referred to textbooks on experimental physical chemistry for details of the construction and use of graphs of this type.

Two other possibilities exist in ternary liquid systems: that in which two components are completely miscible and the third is partially miscible with each, and that in which all combinations of two of the three components are only partially miscible.

Solutions of Gases in Liquids

Nearly all gases are more or less soluble in liquids. One has but to recall the solubility of carbon dioxide, hydrogen sulfide or air in water as common examples.

The amount of gas dissolved in a liquid in general follows *Henry's law*, which states that *the weight of gas dissolved by a given amount of a liquid at a given temperature is proportional to its pressure*. Thus, if the pressure is doubled, twice as much gas will dissolve as at the initial pressure. The extent to which a gas is dissolved in a liquid, at a given temperature, may be expressed in terms of the *solubility coefficient*, which is the volume of gas measured under the conditions of the experiment, that is, absorbed by one volume of the liquid. The degree of solubility also is expressed sometimes in terms of the *absorption coefficient*, which is the volume of gas, reduced to standard conditions, dissolved by one volume of liquid under a pressure of one atmosphere.

Although Henry's law expresses fairly accurately the solubility of slightly soluble gases, it deviates considerably in the case of very soluble gases such as hydrogen chloride and ammonia. Such deviations most frequently are due to chemical interaction of solute and solvent.

The solubility of gases in liquids *decreases* with a *rise in temperature* and, in general, also when salts are added to the solvent, the latter effect being referred to as the *salting-out* of the gas.

Solutions of gases potentially are dangerous when exposed to warm temperatures because of the liberation and expansion of the dissolved gas which may cause the container to burst. Bottles containing such solutions (eg, strong ammonia solution) should be cooled before opening, if practical, and the stopper should be covered with a cloth before attempting its removal.

Thermodynamics of the Solution Process

In the discussions under this heading the solute is assumed to be in the liquid state, hence, the heat of solution ($\Delta H'$) is a term different from that in Eq 3 (ΔH). The heat of solution for a solid solute going into solution as defined in Eq 3 is the net heat effect for the overall dissolution

$$A_{(\text{solid})} \rightleftarrows A_{(\text{liquid})} \rightleftarrows A_{(\text{solution})}$$

Considering only the process

$$A_{(\text{liquid})} \rightleftarrows A_{(\text{solution})}$$

and assuming that the solute is a liquid (or a supercooled liquid in the case of a solid) at a temperature close to room temperature, where the energy needed for melting (heat of fusion) is not being considered.

For a physical or chemical reaction to occur spontaneously at a constant temperature and pressure, the net free-energy change, ΔG, for the reaction should be negative (see *Thermodynamics*, Chapter 15). Furthermore, it is known that the free-energy change depends on heat-related enthalpy ($\Delta H'$) and order-related entropy (ΔS) factors as seen in

$$\Delta G = \Delta H' - T\Delta S \qquad (13)$$

where T is the temperature. Recall, also, that the relation between free energy and the equilibrium constant, K, for a reaction is given by

$$\Delta G = -RT \ln K \qquad (14)$$

Equations 13 and 14 certainly apply to the solution of a drug. Since the solubility is, in reality, an equilibrium constant, Eq 14 indicates that the greater the negative value of ΔG, the greater the solubility.

The interplay of these two factors, $\Delta H'$ and ΔS in Eq 13, determines the free-energy change and, hence, whether or not dissolution of a drug will occur spontaneously. Thus, if in the solution process $\Delta H'$ is negative and ΔS positive, dissolution is favored since ΔG will be negative.

As the heat of solution is quite significant in the dissolution process one must look at its origin. (For an excellent and more complete discussion of the interactions and driving forces underlying the dissolution process, see Higuchi.[4]) The mechanism of solubility involves severing of the bonds that hold together the ions or molecules of a solute, the separation of molecules of solvent to create a space in the solvent into which the solute can be fitted and the ultimate response of solute and solvent to whatever forces of interaction may exist between them. In order to sever the bonds between molecules or ions of solute in the liquid state, energy must be supplied, as is the case also when molecules of solvent are to be separated. If heat is the source of energy it is apparent that both processes require the absorption of heat.

Solute–solvent interaction, on the other hand, generally is accompanied by the evolution of heat since the process occurs spontaneously. In effecting solution there is, accordingly, a heat-absorbing effect and a heat-releasing effect to be considered beyond those required to melt a solid. If there is no, or very little, interaction between solute and solvent, the only effect will be that of absorption of heat to produce the necessary separations of solute and solvent molecules or ions. If there is a significant interaction between solute and solvent, the amount of heat in excess of that required to overcome the solute–solute and the solvent–solvent forces is liberated. If the opposing heat effects are equal, there will be no change of temperature.

When $\Delta H'$ is zero, and there is no volume change, an *ideal solution* is said to exist since the solute–solute, solvent–solvent and solute–solvent interactions are the same. For such an ideal solution, the solubility of a solid can be predicted from its heat of fusion (the energy needed to melt the solid) at temperatures below its melting point. The student is referred to Ref 1, page 281 to see how this calculation is made.

When the heat of solution has a positive (energy absorbed) or negative (energy liberated) value, the solution is said to be a *nonideal solution*. A negative heat of solution favors solubility while a positive heat works against dissolution.

The magnitude of the various attractive forces involved between solute, solvent and solute–solvent molecules may vary greatly and thus could lead to varying degrees of positive or negative enthalpy changes in the solution process. The reason for this is that the molecular structure of the various solutes and solvents determining the interactions can themselves vary greatly. (For a discussion of these effects, see Ref 1, page 41.)

The solute–solute interaction that must be overcome can vary from the strong ion–ion interaction (as in a salt), to the weaker dipole–dipole interaction (as in nearly all organic medicinals that are not salts), to the weakest induced dipole-induced dipole interaction (as with naphthalene).

The attractive forces in the solvent that must be overcome are, most frequently, the dipole–dipole interaction (as found in water or acetone) and the induced dipole–induced dipole interaction (as in liquid petrolatum).

The energy-releasing solute–solvent interactions that must be taken into account may be one of four types. In decreasing energy of interaction these are ion–dipole interactions (eg, a sodium ion interacting with water), dipole–dipole interactions (eg, an organic acid dissolved in water), dipole–induced dipole interaction, to be discussed later (eg, an organic acid dissolved in carbon tetrachloride) and induced dipole–induced dipole interactions (eg, naphthalene dissolved in benzene).

Since the energy-releasing solute–solvent interaction should approximate the energy needed to overcome the solute-solute and solvent–solvent interactions, it should be apparent why it is not possible to dissolve a salt like sodium chloride in benzene. The interaction between the ions and benzene does not supply enough energy to overcome the interaction between the ions in the solute and therefore gives rise to a positive heat of solution. On the other hand, the interaction of sodium and chloride ions with water molecules does provide an amount of energy approximating the energy needed to separate the ions in the solute and the molecules in the solvent.

Consideration must next be given to entropy effects in dissolution processes. Entropy is an indicator of the disorder or randomness of a system. The more positive the entropy change (ΔS) is, the greater the degree of randomness or disorder of the reaction system and the more favorably disposed is the reaction. Unlike $\Delta H'$, the entropy change (an entropy of mixing) in an ideal solution, is not zero, but has some positive value since there is an increase in the disorderliness or entropy of the system upon dissolution. Thus, in an ideal solution with $\Delta H'$ zero and ΔS positive, ΔG would have a negative value and the process, therefore, would be spontaneous.

In a nonideal solution, on the other hand, where $\Delta H'$ is not zero, ΔS can be equal to, greater than or less than the entropy of mixing found for the ideal solution. A nonideal solution with an entropy of mixing equal to that of the ideal solution is called a "regular solution." These solutions usually occur with nonpolar or weakly polar solutes and solvents. Such solutions are accompanied by a positive enthalpy change, implying that the solute–solvent molecular interaction is less than the solute–solute and solvent–solvent molecular interactions. Regular solutions are amenable to rigorous physical chemical analysis which will not be covered

in this chapter but which can be found in outline form in Ref 1, page 282.

The possibility exists in a nonideal solution that the entropy change is greater than for an ideal solution. Such a solution occurs when there is an association among solute or solvent molecules. In essence, then, the dissolution process occurs when one begins at a relatively ordered (low entropy) state and progresses to a disorderly (high entropy) state. The overall entropy change is positive, greater than that of the ideal case, and favorable to dissolution. As may be expected, the enthalpy change in such a solution is positive since association in a solute or solvent must he overcome. The facilitated solubility of citric acid (an unsymmetrical molecule), as compared to inositol (a symmetrical molecule), may be explained on the basis of such a favorable entropy change.[4]

The solubility of citric acid is greater than that of inositol, yet, on the basis of their heats of solution, inositol should be more soluble. One may regard this phenomenon in another way. The reason for the higher solubility of citric acid is that although there is no hindrance in the transfer of a citric acid molecule as it goes from the solute to the solution phase, when the structurally unsymmetrical citric acid attempts to return to the solute phase from solution, it must assume an orientation that will allow ready interaction with polar groups already oriented. If it does not have the required orientation, it will not return readily to the solute, but it will remain in solution, thus bringing about a solubility larger than expected on the basis of heat of solution.

On the other hand, the structurally symmetrical inositol, as it leaves the solution phase, can interact with the solute phase without requiring a definite orientation; all orientations are equivalent. Hence, inositol can enter the solute phase without hindrance and, therefore, no facilitation of its solubility is observed.

In general, unsymmetrical molecules tend to be more soluble than symmetrical molecules.

Another type of nonideal solution occurs where there is an entropy change less than that expected of an ideal solution. Such nonideal behavior can occur with polar solutes and solvents. In a nonideal solution of this type there is significant interaction between solute and solvent. As may be expected, the enthalpy change ($\Delta H'$) in such a solution is negative and favors dissolution, but this effect is tempered by the unfavorable entropy change occurring at the same time. The reason for the lower-than-ideal entropy change can be visualized where the equilibrium system is more orderly and has a lower entropy than that expected for an ideal solution. The overall entropy change of solution, thus, would be less and not favorable to dissolution. One may rationalize the lower-than-expected solubility of lithium fluoride on the basis of this phenomenon. Compared with other alkali halides, it has a solubility lower than would be expected based solely on enthalpy changes. Because of the small size of ions in this salt there may be considerable ordering of water molecules in the solution. This effect must, of course, lead to a lowered entropy and an unfavorable effect on solubility. The effect of soluble salts on the solubility of nonelectrolytes (page 209) or slightly soluble salts (page 210) may be considered a result of an unfavorable entropy effect.

Pharmaceutical Solvents

The discussion will focus now on solvents available to pharmacists and, in particular, on their interactions and properties of these solvents. It is most important that the pharmacist obtain an understanding of the possible differences in solubility of a given solute in various solvents since he most often is called on to select a solvent which will dissolve the solute. A knowledge of the properties of solvents will allow the intelligent selection of suitable solvents.

Molecular Interactions

The solvent–solvent interaction is, in pharmaceutical solvents, always made up of a dipole–dipole interaction (Keesom Force) and an induced dipole–induced dipole interaction (London Force). It is important to keep in mind that both forces are always present; the contribution that each of these forces makes toward the overall attractive force depends on the structure of the solvent molecule. Some solvents have interactions which predominantly involve the Keesom Force (eg, water), while others are predominantly composed of the London Force (eg, chloroform); usually, both forces will be found.

Dipole–Dipole Forces—The unequal sharing of the electron pair between two atoms due to a difference in their electronegativity brings about a separation of the positive and negative centers of electricity in the molecule, causing it to become polarized; that is, to assume a partial ionic character. The molecule then is said to be a *permanent dipole* and the substance described as being a *polar compound*.

The greater the difference in the electronegativities of the constituent atoms, the greater the inequality of sharing of the electron pair, the greater the distance between the positive and negative centers of electricity in the molecule and the more polar the resulting molecule. As the character of the bonds are intermediate between those existing in nonpolar compounds and those occurring in ionic salts, it is to be expected that the properties of polar compounds should be intermediate between those of the two other classes. Such, in fact, generally is the case.

Coordinate covalent compounds all are very strongly polar because both of the electrons constituting the bonding pair have been contributed by a donor atom which, in effect, loses an electron and becomes positively charged, while the acceptor atom may be considered to gain an electron and become negatively charged.

While, in general, the electronegativities of different kinds of atoms are different, and the expectation is, therefore, that all molecules containing two or more different atoms will be polar, many such molecules actually are nonpolar. Thus, while the electronegativity of chlorine is appreciably different from that of carbon, the molecule of carbon tetrachloride, CCl_4, is nonpolar because the symmetrical arrangement of chlorine atoms about the carbon atom is such as to cancel the effects of the difference in the electronegativity of the constituent atoms. The same is true in the case of methane, CH_4, and for hydrocarbons generally. But the molecules CH_3Cl, CH_2Cl_2 and $CHCl_3$ definitely are polar because of the unsymmetrical distribution of the forces within the molecule.

A knowledge of the degree of polarity of various molecules usually is available in the measurement of the *dipole moment*, μ, of the molecules. This quantity is defined as the product of one of the charges on the molecule and the distance between the two average centers of positive and negative electricity. Measurements of the dipole moment of a substance are made, when possible, on the vapor of the substance but, when not possible, a dilute solution of the substance in a nonpolar solvent is employed. Table II lists the values of the dipole moment for a number of substances.

As stated previously, the molecules of nonpolar substances are characterized by weak attractions for one another, while molecules of polar substances exhibit a relatively strong attraction, which is all the more powerful the greater the dipole moment. The reason for this is readily apparent; the dipoles tend to align themselves so that the opposite charges of two different molecules are adjacent. They affect

Table II—Dipole Moments

Substance	Electrostatic Units ($\mu \times 10^{18}$)
Water	1.85
Acetone	2.8
Methyl alcohol	1.68
Ethyl alcohol	1.70
Phenol	1.70
Ethyl ether	1.14
Aniline	1.51
Nitrobenzene	4.19
o-Dinitrobenzene	6.0
m-Dinitrobenzene	3.8
p-Dinitrobenzene	0.3
Benzene	0
Methane	0
Chloromethane	1.86
Dichloromethane	1.58
Chloroform	1.05
Carbon tetrachloride	0
Carbon monoxide	0.11
Carbon dioxide	0
Oxygen	0
Hydrogen	0
Hydrogen chloride	1.03
Hydrogen bromide	0.78
Hydrogen iodide	0.38
Hydrogen sulfide	0.95
Hydrogen cyanide	2.93
Ammonia	1.49

each other in somewhat the same manner as do two bar magnets, the opposite poles of which are adjacent. While thermal agitation tends to break up the alignment or association of the dipoles, there is, nevertheless, a resultant significant intermolecular force present.

Induced Dipole–Induced Dipole Forces—It is of interest to inquire at this point what force does exist between the molecules of compounds which are nonpolar, eg, those which have zero dipole moment. If some attractive force did not exist, the molecules could not be expected to cling together, as in the solid and liquid states. Although the attraction is relatively slight, there is a force that arises from momentary polarization of the molecules because of electronic oscillations which are taking place continuously within the molecules. The *temporary dipoles* thus produced induce opposite polarizations in adjacent molecules and the net effect is that there is a small but definite attractive force between the molecules to keep them together in the liquid and solid states. This attraction resulting from mutual polarization commonly is referred to as the London Force and as an induced dipole–induced dipole force.

The Hydrogen Bond—The attraction between oppositely charged ends of two dipoles is accentuated when the positive end of one dipole contains a hydrogen atom and the negative end of the other dipole contains an atom of fluorine, oxygen or nitrogen. In such instances the nucleus of the hydrogen atom—which is a proton—appears to be able to bind together the negative end of the molecule, of which it is a part, with the negative end of the adjacent molecule. This may be represented by Fig 16-8.

Since the proton is the smallest positively charged atomic particle, it can draw together two negatively charged atoms

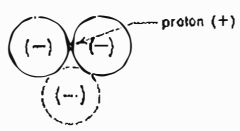
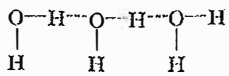


Fig 16-8. Hydrogen bonding.

or ions more closely than can any other—and necessarily larger—positively charged particle. Not more than two negative atoms are capable of being attracted at any given instant, as is evident from Fig 16-8, where a third negative atom is shown to be restricted physically from direct contact with the proton. Water is an excellent example of a substance, the molecules of which are associated through the formation of such a bond—called the *hydrogen bond*. An illustration of such bonding in the case of water may be represented as

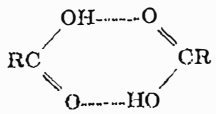


Each dotted line represents the bond or "bridge" established by the hydrogen atom of one water molecule with the oxygen atom of another. It is to be noted that the water molecule is pictured as an angular, rather than as a linear, molecule (H—O—H). This is in accord with the bond angles imposed by the directional character of the bonding orbitals making up the molecule (see Chapter 22). By virtue of its kernel containing six unneutralized protons, not only the valence electrons of the oxygen atom, but also those of the hydrogen atoms are attracted so strongly to the oxygen atom as to make the latter charged negatively, while the rest of the molecule is charged positively.

The hydrogen bond is not a strong bond, but it plays an important role in determining the properties of substances in which it occurs. For example, it primarily is responsible for the unusual properties of water. If the substance H_2O followed the course of the related substances H_2Te, H_2Se and H_2S, in so far as the physical properties of these latter substances are concerned, the freezing point of water would be about −100° and its boiling point about −80°. The unexpectedly high values actually observed are attributed to hydrogen bonding between molecules of water. To break such bonds, as for example in vaporizing water in the form of single H_2O molecules during the process of boiling, more energy is required than would be necessary if the water molecules were not linked by hydrogen bonds.

The molecules of at least the low-molecular-weight alcohols similarly are joined by hydrogen bonds to form a lattice-like structure.

Another example of the manner in which the hydrogen bond functions is seen in the case of carboxylic acids. Such acids usually exist in dimeric form, the two molecules being joined by hydrogen bonding, which may be depicted as



This tendency is so pronounced in the case of acetic acid that even in the vapor state the substance exists in dimeric aggregation.

Classification

On the basis of the forces of interaction occurring in solvents one may broadly classify solvents as one of three types:

1. *Polar solvents*—those made up of strong dipolar molecules having hydrogen bonding (water or hydrogen peroxide).
2. *Semipolar solvents*—those also made up of strong dipolar molecules but which do not form hydrogen bonds (acetone or pentyl alcohol).
3. *Nonpolar solvents*—those made up of molecules having a small or no dipolar character (benzene, vegetable oil or mineral oil).

Naturally, there are many solvents that may fit into more than one of these broad classes; for example, chloroform is a weak dipolar compound but generally is considered nonpo-

lar in character, and glycerin could be considered a polar or semipolar solvent even though it is capable of forming hydrogen bonds.

Types

Water—Water is a unique solvent. Besides being a highly associated liquid, giving rise to its high boiling point, it has another very important property, a high dielectric constant. The dielectric constant (ϵ) indicates the effect that a substance has, when it acts as a medium, on the ease with which two oppositely charged ions may be separated. The higher the dielectric constant of a medium, the easier it is to separate two oppositely charged species in that medium. The dielectric constants of a number of liquids are given in Table III. The values listed are relative to a vacuum which, by definition, has a dielectric constant of unity. According to Coulomb's Law the force of attraction (F) between two oppositely charged ions is

$$F = \frac{Z_1 Z_2}{\epsilon r^2} \tag{15}$$

where Z_1 and Z_2 are the charges on the ions, r is the distance separating the oppositely charged ions and ϵ is the dielectric constant of the medium. Equation 15 indicates that the force of interaction between the oppositely charged ions is proportional inversely to the dielectric constant of the medium. Thus, the interactive force between a sodium and chloride ion in water at a distance r would be $1/80$ that of the same ions in a vacuum separated the same distance. Looking at this example in another way, Coulomb's Law suggests that it is much easier to keep sodium and chloride ions apart in water than in a vacuum. Consider another example: the relative ease with which the ions of sodium chloride may be kept apart the same distance in water, as compared to olive oil, would occur in the ratio of 80/3.1; that is, it is 80/3.1 times easier to keep these ions apart in water than it is in olive oil. The ease of solubilizing salts in solvents like water and glycerin can be explained on the basis of their high dielectric constant. In general, also, the more polar the solvent, the greater its dielectric constant.

There is a very close relationship between dielectric constant and the two types of interactions found in all solvents; that is, the dipole–dipole interaction (Keesom) and the induced dipole–induced dipole interaction (London). The dielectric constant is related to these two forces through a quantity called *total molar polarization*, P, which is a measure of the relative ease with which a charge separation may be made within a molecule. The total molar polarization is given by

Table III—Dielectric Constants (at 20°)

Hydrogen cyanide	116
Water	80
Glycerin	46
Ethylene glycol	41
Methyl alcohol	33
Ethyl alcohol	25
n-Propyl alcohol	22
Acetone	21
Aniline	7.0
Chloroform	5.0
Castor oil	4.6
Ethyl ether	4.3
Octyl alcohol	3.4
Olive oil	3.1
Benzene	2.2
Turpentine oil	2.2
Carbon tetrachloride	2.2
Octane	1.9

$$P = \frac{\epsilon - 1}{\epsilon + 2} \cdot \frac{M}{D} \tag{16}$$

where ϵ is the dielectric constant of the substance, M is the molecular weight and D is the density. (For further details, see Ref 1, page 114.) Total molar polarization is in turn composed of two terms

$$P = P_\alpha + P_\mu = \frac{4}{3}\pi N\alpha + \frac{4}{3}\pi N\left(\frac{\mu^2}{3kT}\right) \tag{17}$$

where $P_\alpha = 4/3\ \pi N\alpha$ is the contribution due to induced polarization (the London contribution), and where

$$P_\mu = \frac{4}{3}\pi N\left(\frac{\mu^2}{3kT}\right)$$

is the contribution due to the permanent dipole (the Keesom contribution), N is Avogadro's Number, α is a constant called the polarizability (related to the induced dipole), μ is the dipole moment, k is the Boltzmann constant (1.38×10^{-16} erg/mole/deg) and T is the absolute temperature. Grouping all constant terms, it is possible to rewrite Eq 17 as

$$P = A + B/T$$

and substituting Eq 16, yields

$$\frac{\epsilon - 1}{\epsilon + 2} \cdot \frac{M}{D} = A + \frac{B}{T} \tag{18}$$

The first term on the right-hand side is the contribution to the dielectric constant of the London dispersions; it is not temperature-dependent. The second term on the right-hand side is the contribution to the dielectric constant of the Keesom dispersions. This latter contribution is temperature-dependent because the contribution from the permanent dipole depends on the dipoles aligning themselves, which tendency is opposed by thermal agitation. Thus, it is apparent from Eq 18 (and from common sense) that as temperature increases, the dielectric constant of dipolar solvents will tend to decrease.

Equation 18 also indicates that solvents which have large dipole moments tend to have large dielectric constants because of the contribution of the P_μ term (Eq 17). Water, which has a very large dielectric constant, is estimated to have $2/3$ of its molecular interaction due to a dipole–dipole interaction and $1/3$ due to the induced dipole–induced dipole interaction. On the other hand, compounds such as benzene, with a dipole moment of zero, will have small dielectric constants since the contribution by the P_μ term will drop out of Eq 18.

There is an important concept that should be considered which has been introduced to pharmaceutical systems.[5] It must be recognized that pharmacists frequently are concerned with dissolving relatively nonpolar drugs in aqueous or mixed polar aqueous solvents. To understand what may be happening in such cases, factors concerned with the entropic effects arising from interactions originating with the nonpolar solutes must be considered. Previously it had been noted that the favorable entropic effect on dissolution was due to the disruption of associations occurring among solute or solvent molecules. Now, consider the effects on solubility due to solute interactions in the solution phase. Since the solutes under discussion are relatively nonpolar, the interactions are of the London Force type or a *hydrophobic association*. This hydrophobic association in aqueous solutions may cause significant structuring of water with a resultant ordered or low-entropy system, unfavorable to solution. As is known, the solution of an essentially nonpolar molecule in water is not a favorable process. It should be stressed that this is due to not only an unfavorable enthalpy change but also an unfavorable entropy change generated by

water-structuring. Such an unfavorable entropy change is quite significant in the solution process. As an example of this effect, the aqueous solubility of a series of alkyl *p*-aminobenzoates shows a ten million-fold decrease in solubility in going from the 1-carbon analog to the 12-carbon analog. These findings demonstrate clearly the considerable effect that hydrophobic associations can have.

Alcohols—*Ethanol*, as a solvent, is next in importance to water. An advantage is that growth of microorganisms does not occur in solutions containing alcohol in a reasonable concentration.

Resins, volatile oils, alkaloids, glycosides, etc are dissolved by alcohol, while many therapeutically inert principles, such as gums, albumin and starch, are insoluble, which makes it more useful as a "selective" solvent. Mixtures of water and alcohol, in proportions varying to suit specific cases, are used extensively. They are often referred to as *hydroalcoholic* solvents.

Glycerin is an excellent solvent, although its range is not as extensive as that of water or alcohol. In higher concentrations it has preservative action. It dissolves the fixed alkalies, a large number of salts, vegetable acids, pepsin, tannin, some active principles of plants, etc, but it also dissolves gums, soluble carbohydrates, starch, etc. It is also of special value as a simple solvent, as in phenol glycerite, or where the major portion of the glycerin simply is added as a preservative and stabilizer of solutions that have been prepared with other solvents (see *Glycerines*, Chapter 84).

Propylene glycol, which has been used widely as a substitute for glycerin, is miscible with water, acetone or chloroform in all proportions. It is soluble in ether and will dissolve many essential oils but is immiscible with fixed oils. It is claimed to be as effective as ethyl alcohol in its power of inhibiting mold growth and fermentation.

Isopropyl alcohol possesses solvent properties similar to those of ethyl alcohol and is used instead of the latter in a number of pharmaceutical manufacturing operations. It has the advantage in that the commonly available product contains not over 1% of water, while ethyl alcohol contains about 5% water, often a disadvantage. Isopropyl alcohol is employed in some liniment and lotion formulations. It cannot be taken internally.

General Properties—Low-molecular-weight and polyhydroxy alcohols form associated structures through hydrogen bonds just as in water. When the carbon-atom content of an alcohol rises above five, generally only monomers then are present in the pure solvent. Although alcohols have high dielectric constants, compared to other types of solvents, they are small compared to water. As has been discussed, the solubility of salts in a solvent should be paralleled by its dielectric constant. That is, as the dielectric constant of a series of solvents increases, the probability of dissolving a salt in the solvent increases. This behavior is observed for the alcohols. Table IV, taken from Higuchi,[4] shows how the solubility of salts follows the dielectric constant of the alcohols.

As mentioned earlier, absolute alcohol rarely is used pharmaceutically. However, hydroalcoholic mixtures such as elixirs and spirits frequently are encountered. A very useful generalization is that the dielectric properties of a mixed solvent, such as water and alcohol, can be approximated as the weighted average of the properties of the pure components. Thus, a mixture of 60% alcohol (by weight) in water should have a dielectric constant approximated by

$$\epsilon_{(\text{mixture})} = 0.6(\epsilon_{(\text{alcohol})}) + 0.4(\epsilon_{(\text{water})})$$

$$\epsilon_{(\text{mixture})} = 0.6(25) + 0.4(80) = 47$$

The dielectric constant of 60% alcohol in water is found experimentally to be 43, which is in close agreement with that just calculated. The dielectric constant of glycerin is 46, close to the 60% alcohol mixture. One would, therefore, expect a salt like sodium chloride to have about the same solubility in glycerin as in 60% alcohol. The solubility of sodium chloride in glycerin is 8.3 g/100 g of solvent and in 60% alcohol about 6.3 g/100 g of solvent. This agreement would be even closer if comparisons were made on a volume rather than weight basis. At least qualitatively it can be said that the solubility of a salt in a solvent or a mixed solvent very closely follows the dielectric constant of the medium or, conversely, that the polarity of mixed solvents is paralleled by their dielectric constant, based on salt solubility.

Table IV—Solubilities of Potassium Iodide and Sodium Chloride in Several Alcohols and Acetone[a]

Solvent	g KI/ 100 g Solvent	g NaCl/ 100 g Solvent
Water	148	35.9
Glycerin	...	8.3 (20°)
Propylene Glycol	50	7.1 (30°)
Methanol	17	1.4
Acetone	2.9	...
Ethanol	1.88	0.065
1-Propanol	0.44	0.0124
2-Propanol	0.18	0.003
1-Butanol	0.20	0.005
1-Pentanol	0.089	0.0018

[a] All measurements are at 25°C unless otherwise indicated.

Although the dielectric constant is useful in interpreting the effect of mixed solvents on salt solubility, it cannot be applied properly to the effect of mixed solvents on the solubility of nonelectrolytes. It was seen earlier that unfavorable entropic effects can occur upon dissolution of relatively nonpolar nonelectrolytes in water. Such an effect due to hydrophobic association considerably affects solubility. Yalkowsky[5] studied the ability of cosolvent systems to increase the solubility of nonelectrolytes in polar solvents where the cosolvent system essentially brings about a reduction in structuring of solvent. Thus, by increasing, in a positive sense, the entropy of solution by using cosolvents, it was possible to increase the solubility of the nonpolar molecule. Using as an example the solubility of alkyl *p*-aminobenzoates in propylene glycol-water systems, Yalkowsky[5] reported that it is possible to increase the solubility of the nonelectrolyte by several orders of magnitude by increasing the fraction of propylene glycol in the aqueous system. Sometimes, it is found that, as a good first approximation, the logarithm of the solubility is related linearly to the fraction of propylene glycol added by

$$\log S_f = \log S_{f=0} + \epsilon f$$

where S_f is the solubility in the mixed aqueous system containing the volume fraction f of nonaqueous cosolvent, $S_{f=0}$ is the solubility in water and ϵ is a constant (not dielectric constant) characteristic of the system under study. Specifically, when a 50% solution of propylene glycol in water is used, there is a 1000-fold increase in solubility of dodecyl *p*-aminobenzoate, in comparison to pure water.

In a series of studies, Martin *et al*[6] have made attempts to predict solubility in mixed solvent systems through an extension of the "regular solution" theory. The equations are logarithmic in nature and can reduce in form to the equations of Yalkowsky.[5]

Acetone and Related Semipolar Materials—Even though acetone has a very high dipole moment (2.8×10^{-18} esu), as a pure solvent it does not form associated structures.

Fig 16-9. The charge separation in acetone.

This is evidenced by its low boiling point (57°) in comparison with the boiling point of the lower-molecular-weight water (100°) and ethanol (79°). The reason why it does not associate is because the positive charge in its dipole does not reside in a hydrogen atom (Fig 16-9), precluding the possibility of its forming a hydrogen bond. However, if some substance which is capable of forming hydrogen bonds, such as water or alcohol, is added to acetone, a very strong interaction through hydrogen bonding will occur (see *Mechanism of Solvent Action* below). Some substances which are semipolar and similar to acetone are aldehydes, low-molecular-weight esters, other ketones and nitro-containing compounds.

Nonpolar Solvents—This class of solvents includes fixed oils such as vegetable oil, petroleum ether (ligroin), carbon tetrachloride, benzene and chloroform. On a relative basis there is a wide range of polarity among these solvents; for example, benzene has no dipole moment while that of chloroform is 1.05×10^{-18} esu. But even the polarity of these compounds normally classified as nonpolar is still in line with the dielectric constant of the solvent. The relation between these quantities is seen best through a quantity called *molar refraction*. The molar refraction (or *refractivity*), R, of a compound is given by

$$R = \frac{n^2 - 1}{n^2 + 2} \cdot \frac{M}{D} \qquad (19)$$

where n is the refractive index of the liquid, M is its molecular weight and D is its density. The similarity between Eq 19 and Eq 16 is to be noted and, indeed, in refractive index measurements using very long wavelengths of light, $n^2 = \epsilon$. Thus, molar refraction under these conditions approximates total molar polarization. Since, in the more nonpolar solvents there is generally no dipole moment, μ, total molar polarization reflects polarization due only to the induced dipoles possible. Thus

$$P_\infty = \frac{n^2 - 1}{n^2 + 2} \cdot \frac{M}{D} = \frac{\epsilon - 1}{\epsilon + 2} \cdot \frac{M}{D} = \frac{4}{3} \pi N \alpha \qquad (20)$$

It is evident from this that the refractive index of a nonpolar compound reflects its relative polarity. For example, the more-polar benzene ($\epsilon = 2.2$) has a higher refractive index, 1.501, than the less-polar hexane ($\epsilon = 1.9$), whose refractive index is 1.375.

It should be emphasized again that when a solvent (such as chloroform) has highly electronegative halogen atoms attached to a carbon atom also containing at least one hydrogen atom, such a solvent will be capable of forming strong hydrogen bonds with solutes which are polar in character. Thus, through the formation of hydrogen bonds such solvents will dissolve polar solutes. For example, it is possible to dissolve alkaloids in chloroform.

Mechanism of Solvent Action

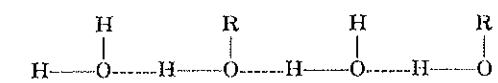
A solvent may function in one, or more, of several ways. When an ionic salt is dissolved, eg, by water, the process of solution involves separation of the cations and anions of the salt with attendant orientation of molecules of the solvent about the ions. Such orientation of solvent molecules about the ions of the solute—a process called *solvation* (*hydration*, if the solvent is water)—is possible only when the solvent is highly polar, whereby, the dipoles of the solvent are attracted to and held by the ions of the solute. The solvent also must possess the ability to keep the solvated, charged ions apart with minimal energy. The role of the dielectric constant in keeping this energy to a minimum has been discussed earlier.

A polar liquid such as water may exhibit solvent action also by virtue of its ability to break a covalent bond in the solute and bring about ionization of the latter. For example, hydrogen chloride dissolves in water and functions as an acid as a result of

$$HCl + H_2O \rightarrow H_3O^+ + Cl^-$$

The ions formed by this preliminary reaction of breaking the covalent bond subsequently are maintained in solution by the same mechanism as ionic salts.

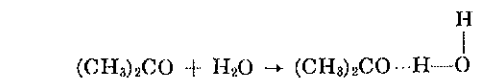
Still another mechanism by which a polar liquid may act as a solvent is that involved when the solvent and solute are capable of being coupled through hydrogen-bond formation. The solubility of the low-molecular-weight alcohols in water, for example, is attributed to the ability of the alcohol molecules to become part of a water-alcohol association complex.



As the molecular weight of the alcohol increases, it becomes progressively less polar and less able to compete with water molecules for a place in the lattice-like arrangement formed through hydrogen bonding; high-molecular-weight alcohols are, therefore, poorly soluble or insoluble in water. When the number of carbon atoms in a normal alcohol reaches five, its solubility in water is reduced materially.

When the number of hydroxyl groups in the alcohol is increased, its solubility in water generally is increased greatly; it is principally, if not entirely, for this reason that such high-molecular-weight compounds as sugars, gums, many glycosides and synthetic compounds, such as the polyethylene glycols, are very soluble in water.

The solubility of ethers, aldehydes, ketones, acids and anhydrides in water, and in other polar solvents, also is attributable largely to the formation of an association complex between solute and solvent by means of the hydrogen bond. The molecules of ethers, aldehydes and ketones, unlike those of alcohols, are not associated themselves, because of the absence of a hydrogen atom which is capahle of forming the characteristic hydrogen bond. Notwithstanding, these substances are more or less polar because of the presence of a strongly electronegative oxygen atom, which is capable of association with water through hydrogen-bond formation. Acetone, for example, dissolves in water, in all likelihood, principally because of the following type of association:



The maximum number of carbon atoms which may be present per molecule possessing a hydrogen-bondable group, while still retaining water solubility, is approximately the same as for the alcohols.

Although nitrogen is less electronegative than oxygen and, thus, tends to form weaker hydrogen bonds, it is observed that amines are at least as soluble as alcohols containing an equivalent chain length. The reason for this is that alcohols form two hydrogen bonds with a net interaction of 12 kcal/mole. Primary amines can form three hydrogen bonds; two amine protons are shared with the oxygens of two water molecules, and the nitrogen accepts one water proton. The net interaction for the primary amine is between 12 and 13 kcal/mole and, hence, it shows an equal or greater solubility compared with corresponding alcohols.

The solvent action of nonpolar liquids involves a somewhat different mechanism. Because they are unable to form dipoles with which to overcome the attractions between ions of an ionic salt, or to break a covalent bond to produce an ionic compound or form association complexes with a solute, nonpolar liquids are incapable of dissolving polar compounds. They only can dissolve, in general, other nonpolar substances in which the bonds between molecules are weak. The forces involved usually are of the induced dipole–induced dipole type. Such is the case when one hydrocarbon is dissolved in another, or an oil or a fat is dissolved in petroleum ether. Sometimes it is observed that a polar substance, such as alcohol, will dissolve in a nonpolar liquid, such as benzene. This apparent exception to the preceding generalization may be explained by the assumption that the alcohol molecule induces a temporary dipole in the benzene molecule which forms an association complex with the solvent molecules. A binding force of this kind is referred to as a *permanent dipole–induced dipole force.*

Some Useful Generalizations—The preceding discussion indicates that enough is known about the mechanism of solubility to be able to formulate some generalizations concerning this important physical property of substances. Because of the greater importance of organic substances in the field of medicinal chemistry, certain of the more useful generalizations with respect to organic chemicals are presented here in summary form. It should be remembered, however, that the phenomenon of solubility usually involves several variables, and there may be exceptions to general rules.

One general maxim which holds true in most instances is: *the greater the structural similarity between solute and solvent, the greater the solubility.* As often stated to the student, "like dissolves like." Thus, phenol is almost insoluble in petroleum ether but is very soluble in glycerin.

Organic compounds containing polar groups capable of forming hydrogen bonds with water are soluble in water, providing that the molecular weight of the compound is not too great. It is demonstrated easily that the polar groups OH, CHO, COH, CHOH, CH_2OH, COOH, NO_2, CO, NH_2 and SO_3H tend to increase the solubility of an organic compound in water. On the other hand, nonpolar or very weak polar groups, such as the various hydrocarbon radicals, reduce solubility; the greater the number of carbon atoms in the radical, the greater the decrease in solubility. Introduction of halogen atoms into a molecule in general tends to decrease solubility because of an increased molecular weight without a proportionate increase in polarity.

Table V—Demonstration of Solubility Rules

Chemical Compound	Solubility[a]
Aniline, $C_6H_5NH_2$	28.6
Benzene, C_6H_6	1430
Benzoic acid, C_6H_5COOH	275
Benzyl alcohol, $C_6H_5CH_2OH$	25
1-Butanol, C_4H_9OH	12
t-Butyl alcohol, $(CH_3)_3COH$	Miscible
Carbon tetrachloride, CCl_4	2000
Chloroform, $CHCl_3$	200
Fumaric acid, *trans*-butenedioic acid	150
Hydroquinone, $C_6H_4(OH)_2$	14
Maleic acid, *cis*-butenedioic acid	5
Phenol, C_6H_5OH	15
Pyrocatechol, $C_6H_4(OH)_2$	2.3
Pyrogallol, $C_6H_3(OH)_3$	1.7
Resorcinol, $C_6H_4(OH)_2$	0.9

[a] The number of mL of water required to dissolve 1 g of solute.

The greater the number of polar groups contained per molecule, the greater the solubility of a compound, provided that the size of the rest of the molecule is not altered; thus, pyrogallol is much more soluble in water than phenol. The *relative positions* of the groups in the molecule also influence solubility; thus, in water, resorcinol (*m*-dihydroxybenzene) is more soluble than catechol (*o*-dihydroxybenzene), and the latter is more soluble than hydroquinone (*p*-dihydroxybenzene).

Polymers and compounds of high molecular weight generally are insoluble or only very slightly soluble.

High melting points frequently are indicative of low solubility for organic compounds. One reason for high melting points is the *association* of molecules and this cohesive force tends to prevent dispersion of the solute in the solvent.

The *cis* form of an isomer is more soluble than the *trans* form. See Table V.

Solvation, which is evidence of the existence of a strong attractive force between solute and solvent, enhances the solubility of the solute, provided there is not a marked ordering of the solvent molecules in the solution phase.

Acids, especially strong acids, usually produce water-soluble salts when reacted with nitrogen-containing organic bases.

Colligative Properties of Solutions

Up to this point concern has been with dissolving a solute in a solvent. Having brought about the dissolution, the solution, quite naturally, has a number of properties which are different from that of the pure solvent. Of very great importance are the *colligative properties* which a solution possesses.

The colligative properties of a solution are those that depend on the *number* of solute particles in solution, irrespective of whether these are molecules or ions, large or small. Ideally, the effect of a solute particle of one species is considered to be the same as that of an entirely different kind of particle, at least in dilute solution. Practically, there may be differences which may become substantial as the concentration of the solution is increased.

The colligative properties which will be considered are:

1. Osmotic pressure.
2. Vapor-pressure lowering.
3. Boiling-point elevation.
4. Freezing-point depression.

Of these four, all of which are related, osmotic pressure has the greatest direct importance in the pharmaceutical sciences. It is the property that largely determines the physiological acceptability of a variety of solutions used for therapeutic purposes.

Osmotic-Pressure Elevation

Diffusion in Liquids—Although the property of diffusion is rapid in gaseous systems, it is not limited to such systems. That molecules or ions in liquid systems possess this same freedom of movement may be demonstrated by placing carefully a layer of water on a concentrated aqueous solution of any salt. In time it will be observed that the boundary between solvent and solution widens gradually since salt moves into the water layer and water migrates from its layer into the salt solution below. Eventually, the composition of the new solution will become uniform throughout. This experiment indicates that *substances tend to move or diffuse from regions of higher concentration to regions of lower concentrations* so that differences in concentration eventually disappear.

Osmosis—In carrying out the experiment just described, it is impossible to distinguish between the diffusion of the solute and that of the solvent. However, by separating the solution and the solvent by means of a membrane that is permeable to the solvent, but not to the solute (such a membrane is referred to as a *semipermeable* membrane), it is possible to demonstrate visibly the diffusion of solvent into the concentrated solution, since volume changes will occur. In a similar manner, if two solutions of different concentra-

tion are separated by a membrane, the solvent will move from the solution of lower solute concentration to the solution of higher solute concentration. This diffusion of solvent through a membrane is called *osmosis*.

There is a difference between the activity or escaping tendency of the water molecules found in the solvent and salt solution separated by the semipermeable membrane. Since *activity*, which is related to water concentration, is higher on the pure solvent side, water moves from solvent to solution in order to equalize escaping-tendency differences. The difference in *escaping-tendency* gives rise to what is referred to as the *osmotic pressure* of the solution, which might be visualized as follows. A semipermeable membrane is placed over the end of a tube and a small amount of salt solution placed over the membrane in the tube. The tube then is immersed in a trough of pure water so that the upper level of the salt solution initially is at the same level as the water in the trough. With time, solvent molecules will move from solvent into the tube. The height of the solution will rise until the *hydrostatic pressure* exerted by the column of solution is equal to the *osmotic pressure*.

Osmotic Pressure of Nonelectrolytes—From quantitative studies with solutions of varying concentration of a solute that does not ionize, it has been demonstrated that *osmotic pressure is proportional to the concentration of the solute*; ie, twice the concentration of a given nonelectrolyte will produce twice the osmotic pressure in a given solvent. (This is not strictly true in solutions of fairly high solute concentration, but does hold quite well for dilute solutions.)

Furthermore, the osmotic pressures of solutions of different nonelectrolytes are proportional to the number of molecules in each solution. Stated in another manner, the osmotic pressures of two nonelectrolyte solutions of the same molal concentration are identical. Thus, a solution containing 34.2 g of sucrose (mol wt 342) in 1000 g of water has the same osmotic pressure as a solution containing 18.0 g of anhydrous dextrose (mol wt 180) in 1000 g of water. These solutions are said to be *isoosmotic* with each other because they have identical osmotic pressures.

A study of the results of osmotic-pressure measurements on different substances led the Dutch chemist Jacobus Henricus van't Hoff, in 1885, to suggest that the solute in a solution may be considered as being analogous to the molecules of a gas and the osmotic pressure as being produced by the bombardment of the semipermeable membrane by the molecules of solute. According to van't Hoff's theory the osmotic pressure of a solution is equal to the pressure which the dissolved substance would exert in the gaseous state if it occupied a volume equal to the volume of the solution. From this it follows that, just as in the case of a gas, there is a proportionality between pressure and concentration of dissolved substance. This proportionality is illustrated well by the values of the osmotic pressure of solutions of sucrose at 0° as determined by the Earl of Berkeley and EGJ Hartley and shown in Table VI.

In column PV of Table VI a quantitative confirmation, at least for fairly dilute solutions, of van't Hoff's oversimplified

Table VI—Osmotic Pressure of Sucrose Solutions

Conc. (g/L), C	Vol in L in Which 1 g Mole is Dissolved, V[a]	Pressure in Atmos P	P/C	PV
10.00	34.2	0.65	0.065	22.2
20.00	17.1	1.27	0.064	21.7
45.00	7.60	2.91	0.065	22.1
93.75	3.65	6.23	0.067	22.7

[a] These figures were obtained by calculating the volume of solution in which 342 (mol wt) g of sucrose would be dissolved.

though useful generalization is shown by the constancy of the values of the product PV. Recall that the product of the pressure and the volume of a gas, at constant temperature, is likewise constant (Boyle's law).

Van't Hoff also deduced that the osmotic pressure must be proportional to the absolute temperature, just as in Charles' law for gases, which deduction was confirmed by the experiments of several workers. From this it follows that the equation $PV = nRT$ is valid for dilute solutions of nonelectrolytes just as a similar equation is valid for gases. However, even as Boyle's law does not apply to gases under high pressures and at low temperatures, so van't Hoff's equation for osmotic pressure does not apply in concentrated solutions.

Osmotic Pressure of Electrolytes—In discussing the generalizations concerning the osmotic pressure of solutions of nonelectrolytes it was stated that the osmotic pressures of two solutions of the same molal concentration are identical. This generalization, however, cannot be made for solutions of electrolytes, ie, acids, alkalies and salts (see Chapter 17).

For example, sodium chloride is assumed to ionize as

$$NaCl \rightarrow Na^+ + Cl^-$$

It is evident that each molecule of sodium chloride that ionizes produces two ions and, if sodium chloride is completely ionized, there will be twice as many particles as would be the case if it were not ionized at all. Furthermore, if each ion has the same effect on osmotic pressure as a molecule, it might be expected that the osmotic pressure of the solution would be twice that of a solution containing the same molal concentration of nonionizing substance.

For solutions which yield more than two ions as, for example

$$K_2SO_4 \rightarrow 2K^+ + SO_4^{2-}$$

$$FeCl_3 \rightarrow Fe^{3+} + 3Cl^-$$

it is to be expected that the complete dissociation of the molecules would give rise to osmotic pressures that are three and four times, respectively, the pressure of solutions containing an equivalent quantity of a nonionized solute. Accordingly, the equation $PV = nRT$, which may be employed to calculate the osmotic pressure of a *dilute* solution of a nonelectrolyte, also may be applied to *dilute* solutions of electrolytes if it is changed to $PV = inRT$, where the value of i approaches the number of ions produced by the ionization of the strong electrolytes cited in the preceding examples. For weak electrolytes i represents the total number of particles, ions and molecules together, in the solution, divided by the number of molecules that would be present if the solute did not ionize. The experimental evidence indicates that in dilute solutions, at least, the osmotic pressures approach the predicted values. It should be emphasized, however, that in more concentrated solutions of electrolytes the deviations from this simple theory are considerable, due to interionic attraction, solvation and other factors.

Biological Aspects of Osmotic Pressure—Osmotic-pressure experiments were made as early as 1884 by the Dutch botanist Hugo de Vries in his study of plasmolysis, which term is applied to the contraction of the contents of plant cells placed in solutions of comparatively high osmotic pressure. The phenomenon is caused by the osmosis of water out of the cell through the practically semipermeable membrane surrounding the protoplasm. If suitable cells (eg, the epidermal cells of the leaf of *Tradescanta discolor*) are placed in a solution of higher osmotic pressure than that of the cell contents, water flows out of the cell, causing the contents to draw away from the cell wall. On the other hand, if the cells are placed in solutions of lower osmotic pressure, water enters the cell, producing an expansion

which is limited by the rigid cell wall. By immersing cells in a series of solutions of varying solute concentration, a solution may be found in which plasmolysis is barely detectable or absent. The osmotic pressure of such a solution is then the same, or very nearly the same, as that of the cell contents, and it is then said that the solution is *isotonic* with the cell contents. Solutions of greater concentration than this are said to be *hypertonic* and solutions of lower concentration, *hypotonic*.

Red blood cells, or erythrocytes, have been studied similarly by immersion into solutions of varying concentration of different solutes. When introduced into water or into sodium chloride solutions containing less than 0.90 g of solute per 100 mL, human erythrocytes swell, and often burst, because of the diffusion of water into the cell and the fact that the cell wall is not sufficiently strong to resist the pressure. This phenomenon is referred to as *hemolysis*. If the cells are placed in solutions containing more than 0.90 g of sodium chloride per 100 mL, they lose water and shrink. By immersing the cells in a solution containing exactly 0.90 g of sodium chloride in 100 mL, no change in the size of the cells is observed; since in this solution the cells maintain their "tone," the solution is said to be *isotonic* with human erythrocytes. For the reasons indicated it is desirable that solutions to be injected into the blood should be made isotonic with erythrocytes. The manner in which this may be done is described in Chapter 79.

Distinction between Isoosmotic and Isotonic—The terms isoosmotic and isotonic are not to be considered as equivalent, although a solution often may be described as being both isoosmotic and isotonic. If a plant or animal cell is in contact with a solution that has the same osmotic pressure as the cell contents, there will be no net gain or loss of water by either solution *provided* the cell membrane is impermeable to all the solutes present. Since the volume of the cell contents remains unchanged, the "tone", or normal state, of the cell is maintained, and the solution in contact with the cell may be described not only as being isoosmotic with the solution in the cell, but also as being isotonic with it. If, however, one or more of the solutes in contact with the membrane can pass through the latter, it is evident that the volume of the cell contents will change, thus altering the "tone" of the cell; in this case the two solutions may be isoosmotic, yet not be isotonic.

It is possible that some substances used in an injection dosage form can cause hemolysis of red blood cells, even when their concentrations are such as to produce solutions theoretically isoosmotic with the cells, because the solutes diffuse through the membrane of the cells. For example, a 1.8% solution of urea has the same osmotic pressure as a 0.9% solution of sodium chloride, but the former solution produces hemolysis of red blood cells; obviously the urea solution is not isotonic with the cells. To determine if a solution is isotonic with erythrocytes, it is necessary to determine the concentration of solute at which the cells retain their normal size and shape. A simple method for doing this was devised by Setnikar and Temelcou,[7] who determined the concentration of a solution at which red blood cells maintained a volume equal to that occupied in an isotonic solution of sodium chloride. The red cell volumes were determined by centrifuging suspensions of them in different solutions, using a hematocrit tube.

Vapor-Pressure Lowering

When a nonvolatile solute is dissolved in a liquid solvent the vapor pressure of the solvent is lowered. This easily can be described qualitatively by visualizing solvent molecules on the surface of the solvent, which normally could escape into the vapor, being replaced by solute molecules which have little if any vapor pressure of their own. For ideal solutions of nonelectrolytes the vapor pressure of the solution follows Raoult's law

$$P_A = X_A P_A^\circ \tag{21}$$

where P_A is the vapor pressure of the solution, P_A° is the vapor pressure of the pure solvent and X_A is the mole fraction of solvent. This relationship states that the vapor pressure of the solution is proportional to the number of molecules of solvent in the solution. Rearranging Eq 21 gives

$$\frac{P_A^\circ - P_A}{P_A^\circ} = (1 - X_A) = X_B \tag{22}$$

where X_B is the mole fraction of the solute. This equation states that the lowering of vapor pressure in the solution relative to the vapor pressure of the pure solvent—called simply the *relative vapor-pressure lowering*—is equal to the mole fraction of the solute. The *absolute* lowering of vapor pressure of the solution is defined by

$$P_A^\circ - P_A = X_B P_A^\circ \tag{23}$$

Example—Calculate the lowering of vapor pressure and the vapor pressure at 20°, of a solution containing 50 g of anhydrous dextrose (mol wt 180.16) in 1000 g of water (mol wt 18.02). The vapor pressure of water at 20°, in absence of air, is 17.535 mm.

First, calculate the lowering of vapor pressure, using Eq 23, in which X_B is the mole fraction of dextrose, defined by

$$X_B = \frac{n_B}{n_A + n_B}$$

where n_A is the number of moles of solvent and n_B is the number of moles of solute. Substituting numerical values

$$n_B = \frac{50}{180.2} = 0.278$$

$$n_B = \frac{1000}{18.02} = 55.5$$

$$X_B = \frac{0.278}{55.5 + 0.278} = 0.00498$$

the lowering of vapor pressure is

$$P_A^\circ - P_A = 0.00498 \times 17.535$$

$$= 0.0873 \text{ mm}$$

The vapor pressure of the solution is

$$P_A = 17.535 - 0.0873$$

$$= 17.448 \text{ mm}$$

Boiling-Point Elevation

In consequence of the fact that the vapor pressure of any solution of a nonvolatile solute is less than that of the solvent, the boiling point of the solution—the temperature at which the vapor pressure is equal to the applied pressure (commonly 760 mm)—must be higher than that of the solvent. This is clearly evident in Fig 16-10.

The relationship between the elevation of boiling point and the concentration of nonvolatile, nonelectrolyte solute may be derived from the Clausius–Clapeyron equation (see Chapter 15), which is

$$\frac{dP}{dT} = \frac{P \cdot \Delta H_{vap}}{RT^2} \tag{24}$$

Replacing the differential expression dP/dT by $\Delta P/\Delta T_b$, where ΔP is the lowering of vapor pressure and ΔT_b is the

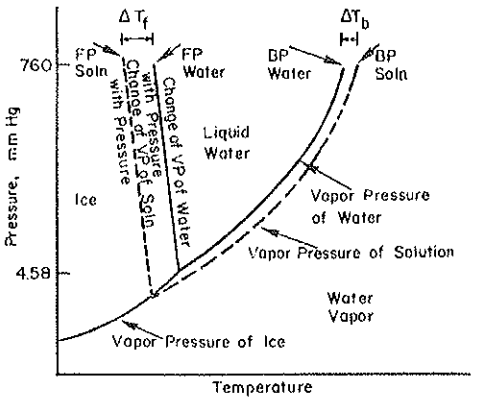


Fig 16-10. Vapor-pressure-temperature diagram for water and an aqueous solution, illustrating elevation of boiling point and lowering of freezing point of the latter.

elevation of boiling point, and introducing $P_A°$, the vapor pressure of the solvent at its boiling point T_0, results in

$$\frac{\Delta P}{\Delta T_b} = \frac{P_A° \cdot \Delta H_{vap}}{RT_0^2} \tag{25}$$

Since the lowering of vapor pressure in an ideal solution is

$$\Delta P = X_B P_A° \tag{26}$$

substitution of this equation into Eq 25, with rearrangement to provide a solution for ΔT_b, gives

$$\Delta T_b = \frac{RT_0^2}{\Delta H_{vap}} X_B \tag{27}$$

This equation may be used to calculate the elevation of the boiling point if the concentration of solute is expressed as the mole fraction. A more common expression, however, is in terms of the molality m (the number of gram-moles of solute per 1000 g of solvent), which relationship is derived as

$$X_B = \frac{n_B}{n_A + n_B} = \frac{m}{1000/M_A + m} \cong \frac{m}{1000/M_A} \tag{28}$$

In these equations M_A is the molecular weight of the solvent. When the solutions are dilute, so that m is small, it may be neglected in the denominator (*but not in the numerator!*) to give the approximate equivalent in Eq 28. Substituting this equivalent into Eq 27 gives

$$\Delta T_b = \frac{RT_0^2 M_A m}{1000 \Delta H_{vap}} \tag{29}$$

Grouping the constants into a single term results in

$$\Delta T_b = K_b m \tag{30}$$

where

$$K_b = \frac{RT_0^2 M_A}{1000 \Delta H_{vap}} \tag{31}$$

and is called the molal boiling-point elevation constant.

The value of this constant for water, which boils at 373.1° K, has a heat of vaporization of 539.7 cal/g and a molecular weight of 18.02, is

$$K_b = \frac{1.987 \times 373.1^2 \times 18.02}{1000 \times 18.02 \times 539.7} = 0.513° \tag{32}$$

Notwithstanding that K_b is called a molal boiling-point elevation constant, it should not be interpreted as the actual rise of boiling point for a 1-molal solution. Such solutions are generally too concentrated to exhibit the ideal behavior assumed in deriving the equation for calculating the theoretical value of the constant. In dilute solutions, however, the actual boiling-point elevation, *calculated to a 1-molal basis*, approaches the theoretical value, the closer the more dilute the solution.

The elevation of boiling point of a dilute solution of a nonelectrolyte solute may be used to calculate the mol wt of the latter. In a solution containing w_B g of solute of M_B in w_A g of solvent the molality m is

$$m = \frac{1000\, w_B}{w_A M_B} \tag{33}$$

substituting this into Eq 30 and rearranging gives

$$M_B = \frac{K_b 1000\, w_B}{w_A \Delta T_b} \tag{34}$$

Freezing-Point Depression

The freezing point of a solvent is defined as the temperature at which the solid and liquid forms of the solvent coexist in equilibrium at a fixed external pressure, commonly 1 atm (760 mm of mercury). At this temperature the solid and liquid forms of the solvent must have the same vapor pressure, for if this were not so, the form having the higher vapor pressure would change into that having the lower vapor pressure.

The freezing point of a solution is the temperature at which the solid form of the pure solvent coexists in equilibrium with the solution at a fixed external pressure, again commonly 1 atm. Since the vapor pressure of a solution is lower than that of its solvent, it is obvious that solid solvent and solution cannot coexist at the same temperature as solid solvent and liquid solvent; only at some lower temperature, where solid solvent and solution do have the same vapor pressure, is equilibrium established. A schematic pressure–temperature diagram for water and an aqueous solution, not drawn to scale and exaggerated for the purpose of more effective illustration, shows the equilibrium conditions involved in both freezing-point depression and boiling-point elevation (Fig 16-10).

The freezing-point lowering of a solution may be quantitatively predicted for ideal solutions, or dilute solutions which obey Raoult's law, by mathematical operations similar to (though somewhat more complex than) those used in deriving the boiling-point elevation constant. The equation for the freezing point lowering, ΔT_f, is

$$\Delta T_f = \frac{RT_0^2 M_A m}{1000 \Delta H_{fus}} = K_f m \tag{35}$$

where

$$K_f = \frac{RT_0^2 M_A}{1000 \Delta H_{fus}} \tag{36}$$

The value of K_f for water, which freezes at 273.1°K and has a heat of fusion of 79.7 cal/g, is

$$K_f = \frac{1.987 \times 273.1^2 \times 18.02}{1000 \times 18.02 \times 79.7} = 1.86° \tag{37}$$

The molal freezing-point depression constant is not intended to represent the freezing-point depression for a 1-molal solution, which is too concentrated for the premise of ideal behavior to be applicable. In dilute solutions the freezing-point depression, calculated to a 1-molal basis, approaches the theoretical value, the agreement between experiment and theory being the better the more dilute the solution.

The freezing point of a dilute solution of a nonelectrolyte solute may be used, as was the boiling point, to calculate the molecular weight of the solute. The applicable equation is

$$M_B = \frac{K_f 1000\, w_B}{w_A \Delta T_f} \tag{38}$$

The molecular weight of organic substances soluble in molten camphor may be determined by observing the freezing point of a mixture of the substance with camphor. This procedure, called the *Rast method*, uses camphor because it has a very large molal freezing-point-depression constant, about 40°. Since the "constant" may vary with different lots of camphor and with variations of technique, the method should be standardized using a solute of known molecular weight.

Freezing-point determinations of molecular weights have the advantage over boiling-point determinations of greater accuracy and precision by virtue of the larger magnitude of the freezing-point depression compared to boiling-point elevation. Thus, in the case of water the molal freezing-point depression is approximately 3.5 times greater than the molal boiling-point elevation.

Relationship between Osmotic Pressure and Vapor-Pressure Depression

The lowering of vapor pressure and the development of osmotic pressure in a solution are both manifestations of the basic condition that the free energy of solvent molecules in the pure solvent is greater than the free energy of solvent molecules in the solution. Consequently, solvent molecules will transfer spontaneously, if given an opportunity, from solvent to solution until equilibrium conditions are established. The transfer can take place either through a membrane permeable only to solvent molecules or, if such contact between solvent and solution is not available, by distillation of solvent from pure solvent to solution, if access through a vapor phase is provided.

If an experiment is performed with two sets of vessels containing solution and solvent, as illustrated in Fig 16-11, differing only in that the long tube of one set has a semipermeable membrane attached to its lower end, while in the other a hypothetical membrane separates the vapor phases, in time the same hydrostatic pressure should develop in both cases. For a definite volume, eg, a mole, of solvent transferred to the solution by distillation the change of free energy, ΔG, in the process is

$$\Delta G = RT \ln \frac{P_A}{P_A^\circ} \tag{39}$$

where P_A is the vapor pressure of the solution and P_A° is the vapor pressure of the solvent.

For the transfer of the same volume of solvent by osmosis the free-energy change is

$$\Delta G = -\bar{V}_A \pi \tag{40}$$

where $\bar{V}_A$ is the partial molal volume of solvent (the volume of 1 mole of solvent in the solution) and π is the osmotic pressure of the solution. Since the free-energy change is the same in both processes

$$-V_A \pi = RT \ln \frac{P_A}{P_A^\circ} \tag{41}$$

rearranging the equation yields

$$\pi = \frac{RT}{\bar{V}_A} \ln \frac{P_A^\circ}{P_A} \tag{42}$$

With this equation the osmotic pressure of a solution may be calculated if its vapor pressure and the partial molal volume of the solvent are known, not only when the solution is sufficiently dilute that Raoult's law is obeyed but also when the concentration is so high as to introduce substantial deviation from the law.

From Eq 42, which has some resemblance to van't Hoff's empirical equation $\pi V = nRT$ for dilute solutions, the latter equation may be derived as follows. If a solution is sufficiently dilute to correspond to Raoult's law, then

$$P_A = X_A P_A^\circ = (1 - X_B) P_A^\circ \tag{43}$$

and then Eq 42 may be written

$$\pi = -\frac{RT}{\bar{V}_A} \ln (1 - X_B) \tag{44}$$

When X_B is small (as in a dilute solution), the term $-\ln (1 - X_B)$ can be shown to be approximately equal to X_B, so that

$$\pi = \frac{RT}{\bar{V}_A} X_B \tag{45}$$

In dilute solutions the approximations $X_B = n_B/n_A$ (where n_B and n_A are the moles of solute and solvent, respectively) and $\bar{V}_A = V/n_A$ (where V is the volume of solution) may be introduced, yielding

$$\pi V = n_B RT \tag{46}$$

which is van't Hoff's equation.

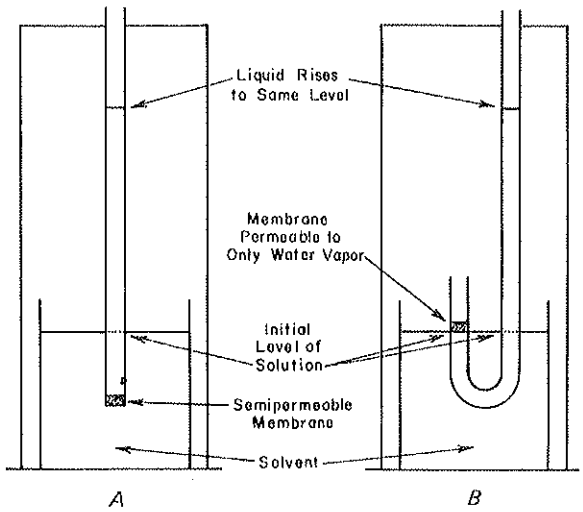


Fig 16-11. Transfer of solvent to equal volumes of solution. *A*: Osmotically, through a semipermeable membrane separating solution and solvent. *B*: By distillation, through a membrane separating solution and solvent vapor.

Ideal Behavior and Deviations

In setting out to derive mathematical expressions for colligative properties, such phrases as "for ideal solutions" or "for dilute solutions" were used to indicate the limitations of the expressions. Samuel Glasstone defines an ideal solution as "one which obeys Raoult's law over the whole range of concentration and at all temperatures" and gives as specific characteristics of such solutions their formation only from constituents which mix in the liquid state without heat change and without volume change. These characteristics reflect the fact that addition of a solute to a solvent produces no change in the forces between molecules of the solvent.

Thus, the molecules have the same *escaping-tendency* in the solution as in the pure solvent and the vapor pressure above the solution is proportional to the ratio of the number of solvent molecules in the surface of the solution to the number of the molecules in the surface of the solvent—which is the basis for Raoult's law.

Any change in intermolecular forces produced by mixing the components of a solution may result in deviation from ideality; such a deviation may be expected particularly in solutions containing both a polar and a nonpolar substance. Solutions of electrolytes, except at high dilution, are especially prone to depart from ideal behavior, even though allowance is made for the additional particles that result from ionization. When solute and solvent combine to form solvates, the escaping-tendency of the solvent may be reduced in consequence of the reduction in the number of free molecules of solvent; thus, a negative deviation from Raoult's law is introduced. On the other hand, the escaping-tendency of the solvent, in a solution of nonvolatile solute, may be increased because the cohesive forces between molecules of solvent are reduced by the solute; this results in a positive deviation from Raoult's law. Chapter 17 considers deviations from ideality in more detail.

While few solutions exhibit ideal behavior over a wide range of concentration, most solutions behave ideally at least in high dilution, where deviations from Raoult's law are negligible.

Comparison of Colligative Properties—In view of the established interrelationships of the colligative properties of ideal solutions or very dilute real solutions, it is possible to predict, by calculation, the magnitude of all these properties of such solutions if the concentration of the nonelectrolyte solute is given. Also, if the magnitude of one of the properties, eg, the freezing point, is known for a solution of unspecified concentration, it is possible to calculate the vapor pressure, boiling point and osmotic pressure, provided the solution is ideal or sufficiently dilute to show negligible deviation from ideality. To what upper limit of concentration a nonideal solution remains "sufficiently dilute" to show ideal behavior is difficult to specify. The answer depends at least in part on the degree of agreement expected between experimental and theoretical values. Certainly, a 1-molal concentration is much too concentrated for a nonideal solution to show conformance with ideal behavior and even in 0.1-molal concentration, deviations are significant and for some purposes may be excessive.

In dealing with colligative properties of solutions which do not behave ideally, caution should be exercised in attempting to predict the magnitude of other colligative properties from one that has been determined experimentally. Earlier, an equation was derived for calculating the vapor pressure of a solution from its osmotic pressure, or *vice versa*, this equation being valid even with solutions showing substantial departures from ideal behavior. The equation is limited, however, to a comparison of these colligative properties at the same temperature. The degree of deviation from ideal behavior for one colligative property will be exactly the same for another only when the temperature is the same for both. It does not follow that the degree of deviation of the colligative properties of a given nonideal solution will be the same for all the properties since at least two of these (freezing point and boiling point) must be determined at quite different temperatures. While in dilute solutions the intermolecular (and/or interionic) forces and interactions *may* change little over the temperature interval between freezing and boiling, in concentrated solutions the change *may* be marked. In the absence of adequate knowledge about the forces and interactions involved, only by experiment can one establish the magnitude of the colligative properties of other than very dilute nonideal solutions. It is important to keep this in mind in estimating the osmotic pressure of a nonideal solution at body temperature from a freezing point determined some 37° lower. While in many cases—possibly the majority of them—such an estimate is warranted by virtue of essential constancy of the various forces and interactions over a wide range of temperature, this is not always the case and the estimate may be significantly inaccurate.

Colligative Properties of Electrolyte Solutions [See Chapter 17]—Earlier in this chapter attention was directed to the increased osmotic pressure observed in solutions of electrolytes, the enhanced effect being attributed to the presence of ions, each of which acts, in general, in the same way as a molecule in developing osmotic pressure. Similar magnification of vapor-pressure lowering, boiling-point elevation and freezing-point depression occurs in solutions of electrolytes. Thus, at a given constant temperature the abnormal effect of an electrolyte on osmotic pressure is paralleled by abnormal lowering of vapor pressure; the other colligative properties are, subject to variation of effect with temperature, comparably intensified. In general the magnitude of each colligative property is proportional to the total number of particles (molecules and/or ions) in solution.

While in *very* dilute solutions the osmotic pressure, vapor-pressure lowering, boiling-point increase and freezing-point depression of solutions of electrolytes would approach values 2, 3, 4, etc times greater for NaCl, Na_2SO_4 and Na_3PO_4 than in solutions of the same molality of a nonelectrolyte, two other effects are observed as the concentration of electrolyte is increased. The first effect results in *less* than 2-, 3- or 4-fold intensification of a colligative property. This reduction is ascribed to interionic attraction between the positively and negatively charged ions, in consequence of which the ions are not completely dissociated from each other and do not exert their full effect in lowering vapor pressure, etc. This deviation generally increases with increasing concentration of electrolyte. The second effect intensifies the colligative properties and is attributed to the attraction of ions for solvent molecules, which holds the solvent in solution and reduces its escaping-tendency, with consequent enhancement of the vapor-pressure lowering. Solvation also may reduce interionic attraction and thereby further lower the vapor pressure. These factors (and possibly others) combine to effect a progressive reduction in the molal values of colligative properties as the concentration of electrolyte is increased to 0.5 to 1 molal, beyond which the molal quantities either increase, sometimes quite abruptly, or remain almost constant.

Activity and Activity Coefficient

Various mathematical expressions are employed to relate properties of chemical systems (equilibrium constants, colligative properties, pH, etc) to the stoichiometric concentration of one or more molecular, atomic or ionic species. In deriving such expressions it either is stated or implied that they are valid only so long as intermolecular, interatomic and/or interionic forces may be ignored or remain constant, under which restriction the system may be expected to behave ideally. But intermolecular, interatomic and/or interionic forces do exist, and not only do they change as a result of chemical reaction but also with variation in the concentration or pressure of the molecules, atoms or ions under observation. In consequence, mathematical expressions involving stoichiometric concentrations or pressures generally have limited applicability. The conventional concentration terms, while providing a count of molecules, atoms or ions per unit volume, afford no indication of the physical or chemical activity of the species measured, and it is this activity which determines the physical and chemical properties of the system.

In recognition of this, GN Lewis introduced both the quantitative concept and methods for evaluation of activity as a true measure of the physical or chemical activity of molecular, atomic or ionic species, whether in the state of gas, liquid or solid, or whether present as a single species or in a mixture. Activity may be considered loosely as a corrected concentration or pressure which takes into account not only the stoichiometric concentration or pressure but also any intermolecular attractions, repulsions or interactions between solute and solvent in solution, association and ionization. Thus, activity measures the net effectiveness of a chemical species. Because only relative values of activity may be determined, a *standard state* must be chosen for quantitative comparisons to be made. Indeed, because activity measurements are needed for many different types of systems, several standard states must be selected. Since this discussion is concerned mainly with solutions, the standard state for the solvent is pure solvent, while for the solute it is a hypothetical solution with free energy corresponding to unit molality under conditions of ideal behavior of the solution. The relationship of activity to concentration is measured in terms of an activity coefficient which is discussed in Chapter 17.

Practical Applications of Colligative Properties

One of the most important pharmaceutical applications of colligative properties is in the preparation of isotonic intravenous and isotonic lacrimal solutions, the details of which are discussed in Chapter 79.

Other applications of the colligative properties are found in experimental physiology. One such application is in the immersion of tissues in salt solutions, which are isotonic with the fluids of the tissue, in order to prevent changes or injuries that may arise from osmosis.

The colligative properties of solutions also may be used in determining the molecular weight of solutes or, in the case of electrolytes, the extent of ionization. The method of determining molecular weight depends on the fact that each of the colligative properties is altered by a constant value when a definite number of molecules of solute is added to a solvent [See Chapter 17]. For example, in dilute solutions the freezing point of water is lowered at the rate of 1.855° for each gram-molecular weight of a nonelectrolyte dissolved in 1000 g of water.*

The constant 1.855° is commonly called the *molal freezing-point depression* of water. To find the molecular weight of a nonelectrolyte, therefore, all that is necessary is to determine the freezing point of a dilute aqueous solution of known concentration of the nonelectrolyte and, by proportion, to calculate the quantity that would produce, theoretically, a depression of 1.855° when 1000 g of water is used as the solvent. If the substance is insoluble in water, it may be dissolved in another solvent, in which case, however, the freezing-point depression of a solution corresponding to a gram-molecular weight of the solute in 1000 g of solvent will be some value other than 1.855°. In the case of benzene, for example, this value is 5.12°; carbon tetrachloride, 2.98°; phenol, 7.27° and camphor, about 40° (see *Freezing-Point Depression*, in this chapter).

The boiling-point elevation may be used similarly for determining molecular weights. The boiling point of water is raised at the rate of 0.52° for each gram-molecular weight of solute dissolved in 1000 g of water,* the corresponding values for benzene, carbon tetrachloride and phenol being 2.57°, 4.88° and 3.60°, respectively. The observation of vapor-pressure lowering and osmotic pressure likewise may be used to calculate molecular weights.

To determine the extent to which an electrolyte is ionized, it is necessary to know its molecular weight, as determined by some other method, and then to measure one of the four colligative properties. The deviation of the results from similar values for nonelectrolytes then is used in calculating the extent of ionization.

References

1. Martin AN *et al:* *Physical Pharmacy*, Lea & Febiger, Philadelphia, 1983.
2. Kramer SF, Flynn GL: *J Pharm Sci 61:* 1896, 1972.
3. Campbell AN, Campbell AJR: *J Am Chem Soc 59:* 2481, 1937.
4. Higuchi T, in Lyman R: *Pharmaceutical Compounding and Dispensing*, Lippincott, Philadelphia, 159, 1949.
5. Yalkowsky, SH: *Techniques of Solubilization of Drugs*, Marcel Dekker, New York, 91, 1981.
6. Martin A *et al: J Pharm Sci, 71:* 849, 1982.
7. Setnikar I, Temelcou O: *JAPhA Sci Ed 48:* 628, 1959.

* These constants apply only to solutions that are considerably more dilute than 1 molal; a substantial deviation would be observed if a 1-molal solution were to be used.

CHAPTER 17

Ionic Solutions and Electrolytic Equilibria

Paul J Niebergall, PhD
Professor of Pharmaceutical Sciences and Director
Pharmaceutical Development Center
Medical University of South Carolina
Charleston, SC 29425

Electrolytes

In the preceding chapter, attention was directed to the colligative properties of nonelectrolytes, or substances whose aqueous solutions do not conduct electricity. Substances whose aqueous solutions conduct electricity are known as electrolytes, and are typified by inorganic acids, bases and salts. In addition to the property of electrical conductivity, solutions of electrolytes exhibit anomalous colligative properties.

Colligative Properties

In general, for nonelectrolytes, a given colligative property of two equimolal solutions will be identical. This generalization, however, cannot be made for solutions of electrolytes.

Van't Hoff pointed out that the osmotic pressure of a solution of an electrolyte is considerably greater than the osmotic pressure of a solution of a nonelectrolyte of the same molal concentration. This anomaly remained unexplained until 1887 when Arrhenius proposed a hypothesis which forms the basis for our modern theories of electrolyte solutions.

This theory postulated that when electrolytes are dissolved in water they split up into charged particles known as ions. Each of these ions carries one or more electrical charges, with the total charge on the positive ions (cations) being equal to the total charge on the negative ions (anions). Thus, although a solution may contain charged particles, it remains neutral. The increased osmotic pressure of such solutions is due to the increased number of particles formed in the process of ionization. For example, sodium chloride is assumed to dissociate as

$$Na^+Cl^- \xrightarrow{H_2O} Na^+ + Cl^-$$

It is evident that each molecule of sodium chloride which is dissociated produces two ions and, if dissociation is complete, there will he twice as many particles as would be the case if it were not dissociated at all. Furthermore, if each ion has the same effect on osmotic pressure as a molecule, it might be expected that the osmotic pressure of the solution would be twice that of a solution containing the same molal concentration of a nonionizing solute.

Osmotic-pressure data indicate that, in very dilute solutions of salts which yield two ions, the pressure is very nearly double that of solutions of equimolal concentrations of nonelectrolytes. Similar magnification of vapor-pressure lowering, boiling-point elevation and freezing-point depression occurs in dilute solutions of electrolytes.

Van't Hoff defined a factor i as the ratio of the colligative effect produced hy a concentration m of electrolyte, divided hy the effect observed for the same concentration of nonelectrolyte, or

$$i = \frac{\pi}{(\pi)_0} = \frac{\Delta P}{(\Delta P)_0} = \frac{\Delta T_b}{(\Delta T_b)_0} = \frac{\Delta T_f}{(\Delta T_f)_0} \qquad (1)$$

in which π, ΔP, ΔT_b and ΔT_f refer to the osmotic pressure, vapor-pressure lowering, boiling-point elevation and freezing-point depression, respectively, of the electrolyte. The terms $(\pi)_0$, etc refer to the nonelectrolyte of the same concentration. In general, with strong electrolytes (those assumed to be 100% ionized), the van't Hoff factor is equal to the numher of ions produced when the electrolyte goes into solution (2 for NaCl and $MgSO_4$, 3 for $CaCl_2$ and Na_2SO_4, 4 for $FeCl_3$ and Na_3PO_4, etc).

In *very* dilute solutions the osmotic pressure, vapor-pressure lowering, boiling-point elevation and freezing-point depression of solutions of electrolytes approach values 2, 3, 4 or more times greater (depending on the type of strong electrolyte) than in solutions of the same molality of nonelectrolyte, thus confirming the hypothesis that an ion has the same primary effect as a molecule on colligative properties. It bears repeating, however, that two other effects are observed as the concentration of electrolyte is increased.

The first effect results in less than 2-, 3- or 4-fold intensification of a colligative property. This reduction is ascribed to interionic attraction between the positively and negatively charged ions, in consequence of which the ions are not dissociated completely from each other and do not exert their full effect on vapor pressure and other colligative properties. This deviation generally increases with increasing concentration of electrolyte.

The second effect intensifies the colligative properties and is attributed to the attraction of ions for solvent molecules (called solvation, or, if water is the solvent, hydration), which holds the solvent in solution and reduces its escaping tendency, with a consequent enhancement of the vapor-pressure lowering. Solvation also reduces interionic attraction and, thereby, further lowers the vapor pressure.

Conductivity

The ahility of metals to conduct an electric current results from the mobility of electrons in the metals. This type of conductivity is called *metallic* conductance. On the other hand, various chemical compounds—notahly acids, bases and salts—conduct electricity by virtue of ions present or formed, rather than by electrons. This is called *electrolytic* conductance, and the conducting compounds are electrolytes. While the fact that certain electrolytes conduct electricity in the molten state is important, their behavior when dissolved in a solvent, particularly in water, is of greater concern in pharmaceutical science.

The electrical conductivity (or conductance) of a solution of an electrolyte is merely the reciprocal of the resistance of the solution. Hence, to measure conductivity is actually to

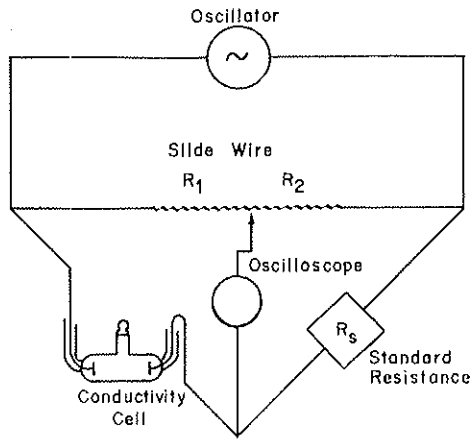


Fig 17-1. Alternating current Wheatstone bridge for measuring conductivity.

measure electrical resistance, commonly with a Wheatstone-bridge apparatus, and then to *calculate* the conductivity. Fig 17-1 is a representation of the component parts of the apparatus.

The solution to be measured is placed in a glass or quartz cell having two inert electrodes, commonly made of platinum or gold and coated with spongy platinum to absorb gases, across which passes an alternating current generated by an oscillator at a frequency of about 1000 Hz. The reason for using alternating current is to reverse the electrolysis that occurs during flow of current and which would cause polarization of the electrodes and lead to abnormal results. The size of the electrodes and their distance apart may be varied to reduce very high resistance and increase very low resistance in order to increase the accuracy and precision of measurement. Thus, solutions of high conductance (low resistance) are measured in cells having small electrodes relatively far apart, while solutions of low conductance (high resistance) are measured in cells with large electrodes placed close to each other.

Electrolytic resistance, like metallic resistance, varies directly with the length of the conducting medium and inversely with its cross-sectional area. The known resistance required for the circuit is provided by a resistance box containing calibrated coils. Balancing of the bridge may be achieved by sliding a contact over a wire of uniform resistance until no (or minimum) current flows through the circuit, as detected either visually with a cathode-ray oscilloscope or audibly with earphones.

The resistance, in ohms, is calculated by the simple procedure used in the Wheatstone-bridge method. The reciprocal of the resistance is the conductivity, the units of which are *reciprocal ohms* (also called *mhos*). As the numerical value of the conductivity will vary with the dimensions of the conductance cell, the value must be calculated in terms of *specific conductance*, L, which is the conductance in a cell having electrodes of 1-sq cm cross-sectional area and 1 cm apart. If the dimensions of the cell used in the experiment were known, it would be possible to calculate the specific conductance, but this information actually is not required, because it is possible—and much more convenient—to calibrate a cell by measuring in it the conductivity of a standard solution of known specific conductance and then calculating a "cell constant." Since this constant is a function only of the dimensions of the cell, it can be used to convert all measurements in that cell to specific conductivity. Solutions of known concentration of pure potassium chloride are used as standard solutions for this purpose.

Equivalent Conductance—In studying the variation of conductance of electrolytes with dilution it is essential to make allowance for the degree of dilution in order that the comparison of conductances may be made for identical amounts of solute. This may be achieved by expressing conductance measurements in terms of *equivalent conductance*, Λ, which is obtained by multiplying the specific conductance, L, by the volume in milliliters, V_e, of solution containing 1 g-eq of solute. Thus

$$\Lambda = LV_e = \frac{1000L}{C} \qquad (2)$$

where C is the concentration of electrolyte in the solution in g-Eq/L, ie, the normality of the solution. For example, the equivalent conductance of 0.01 N potassium chloride solution, which has a specific conductance of 0.001413 mho/cm may be calculated in either of the following ways

$$\Lambda = 0.001413 \times 100{,}000 = 141.3 \text{ mho cm}^2\text{/eq}$$

or

$$\Lambda = \frac{1000 \times 0.001413}{0.01} = 141.3$$

Strong and Weak Electrolytes—It is customary to classify electrolytes broadly as *strong electrolytes* and *weak electrolytes*. The former category includes solutions of strong acids, strong bases and most salts; the latter includes weak acids and bases, primarily organic acids, amines and a few salts. The usual criterion for distinguishing between strong and weak electrolytes is the extent of ionization. An electrolyte existing entirely or very largely as ions is considered a strong electrolyte, while one that is a mixture of a substantial proportion of molecular species along with ions derived therefrom is a weak electrolyte. For the purposes of this discussion, classification of electrolytes as strong or weak will be on the basis of certain conductance characteristics exhibited in aqueous solution.

The equivalent conductances of a number of electrolytes, at different concentrations, are given in Table I and for certain of these electrolytes again in Fig 17-2, where the

Table I—Equivalent Conductances at 25°

g-Eq/L	HCl	HAc	NaCl	KCl	NaI	KI	NaAc
Inf dil	426.1	390.6[a]	126.5	149.9	126.9	150.3	91.0
0.0005	422.7	67.7	124.5	147.8	125.4	...	89.2
0.001	421.4	49.2	123.7	146.9	124.3	...	88.5
0.005	415.8	22.9	120.6	143.5	121.3	144.4	85.7
0.01	412.0	16.3	118.5	141.3	119.2	142.2	83.8
0.02	407.2	11.6	115.8	138.3	116.7	139.5	81.2
0.05	399.1	7.4	111.1	133.4	112.8	135.0	76.9
0.1	391.3	5.2	106.7	129.0	108.8	131.1	72.8

[a] The equivalent conductance at infinite dilution for acetic acid, a weak electrolyte, is obtained by adding the equivalent conductances of hydrochloric acid and sodium acetate and subtracting that of sodium chloride (see text for explanation).

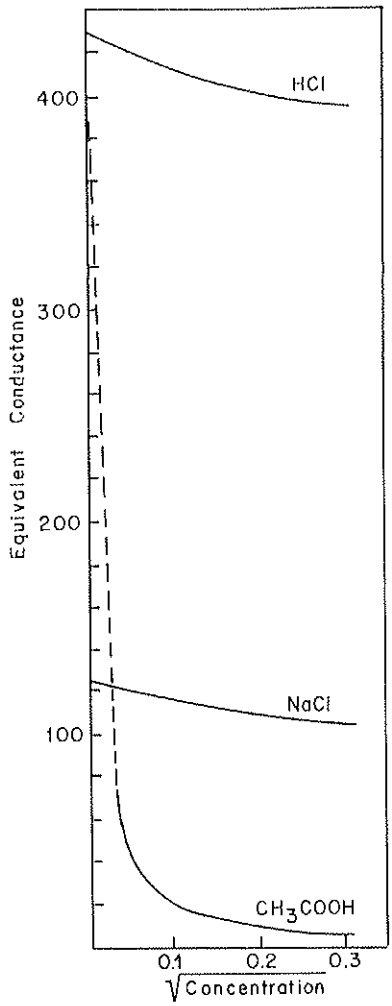


Fig 17-2. Variation of equivalent conductance with square root of concentration.

equivalent conductance is plotted against the square root of concentration. By plotting the data in this manner a linear relationship is observed for strong electrolytes, while a steeply rising curve is noted for weak electrolytes; this difference is a characteristic which distinguishes strong and weak electrolytes. The interpretation of the steep rise in the equivalent conductance of weak electrolytes is that the degree of ionization increases with dilution, becoming complete at infinite dilution. Interionic interference effects generally have a minor role in the conductivity of weak electrolytes. With strong electrolytes, which usually are completely ionized, the increase in equivalent conductance results not from increased ionization but rather from diminished ionic interference as the solution is diluted, in consequence of which ions have greater freedom of mobility, ie, increased conductance.

The value of the equivalent conductance extrapolated to infinite dilution (zero concentration), designated by the symbol Λ_0, has a special significance. It represents the equivalent conductance of the completely ionized electrolyte when the ions are so far apart that there is no interference with their migration due to interionic interactions. It has been shown, by Kohlrausch, that the equivalent conductance of an electrolyte at infinite dilution is the sum of the equivalent conductances of its component ions at infinite dilution, expressed symbolically as

Table II—Equivalent Ionic Conductivities at Infinite Dilution, at 25°

Cations	l_0	Anions	l_0
H^+	349.8	OH^-	198.0
Li^+	38.7	Cl^-	76.3
Na^+	50.1	Br^-	78.4
K^+	73.5	I^-	76.8
NH_4^+	61.9	Ac^-	40.9
$\frac{1}{2}Ca^{2+}$	59.5	$\frac{1}{2}SO_4^{2-}$	79.8
$\frac{1}{2}Mg^{2+}$	53.0		

$$\Lambda_0 = l_0(\text{cation}) + l_0(\text{anion}) \qquad (3)$$

The significance of Kohlrausch's law is that each ion, at infinite dilution, has a characteristic value of conductance that is independent of the conductance of the oppositely charged ion with which it is associated. Thus, if the equivalent conductances of various ions are known, the conductance of any electrolyte may be calculated simply by adding the appropriate ionic conductances. Since the fraction of current carried by cations (*transference number* of the cations) and by anions (transference number of anions) in an electrolyte may be determined readily by experiment, ionic conductances are known. Table II gives the equivalent ionic conductances at infinite dilution of some cations and anions. It is not necessary to have this information in order to calculate the equivalent conductance of an electrolyte, for Kohlrausch's law permits the latter to be calculated by adding and subtracting values of Λ_0 for appropriate electrolytes. For example, the value of Λ_0 for acetic acid may be calculated as

$$\Lambda_0(CH_3COOH) = \Lambda_0(HCl) + \Lambda_0(CH_3COONa) - \Lambda_0(NaCl)$$

which is equivalent to

$$l_0(H^+) + l_0(CH_3COO^-) = l_0(H^+) + l_0(Cl^-) + l_0(Na^+) + l_0(CH_3COO^-) - l_0(Na^+) - l_0(Cl^-)$$

This method is especially useful for calculating Λ_0 for weak electrolytes such as acetic acid. As is evident from Fig 17-2, the Λ_0 value for acetic acid cannot be determined accurately by extrapolation because of the steep rise of conductance in dilute solutions. For strong electrolytes, on the other hand, the extrapolation can be made very accurately. Thus, in the example above, the values of Λ_0 for HCl, CH_3COONa and NaCl are determined easily by extrapolation since the substances are strong electrolytes. Substitution of these extrapolated values, as given in Table II, yields a value of 390.6 for the value of Λ_0 for CH_3COOH.

Ionization of Weak Electrolytes—When Arrhenius introduced his theory of ionization he proposed that the degree of ionization, α, of an electrolyte is measured by the ratio

$$\alpha = \Lambda/\Lambda_0 \qquad (4)$$

where Λ is the equivalent conductance of the electrolyte at any specified concentration of solution and Λ_0 is the equivalent conductance at infinite dilution. As strong electrolytes were not then recognized as being 100% ionized, and interionic interference effects had not been evaluated, he believed the equation to be applicable to both strong and weak electrolytes. Since it now is known that the apparent variation of ionization of strong electrolytes arises from a change in the mobility of ions at different concentrations, rather than from varying ionization, the equation is not applicable to strong electrolytes. It does provide, however, a generally acceptable approximation of the degree of ionization of weak

electrolytes, for which deviations resulting from neglect of activity coefficients and of some change of ionic mohilities with concentration are, for most purposes, negligible. The following example illustrates the use of the equation to calculate the degree of ionization of a typical weak electrolyte.

Example—Calculate the degree of ionization of 1×10^{-3} *N* acetic acid, the equivalent conductance of which is 48.15 mho cm²/Eq. The equivalent conductance at infinite dilution is 390.6 mho cm²/Eq.

$$\alpha = \frac{48.15}{390.6} = 0.12$$

$$\% \text{ ionization} = 100\alpha = 12\%$$

The degree of dissociation also can he calculated using the van't Hoff factor, i, and the following equation

$$\alpha = \frac{i - 1}{\upsilon - 1} \tag{5}$$

where υ is the number of ions into which the electrolyte dissociates.

Example—A 1.0×10^{-3} *N* solution of acetic acid has a van't Hoff factor equal to 1.12. Calculate the degree of dissociation of the acid at this concentration.

$$\alpha = \frac{i - 1}{\upsilon - 1} = \frac{1.12 - 1}{2 - 1} = 0.12$$

This result agrees with that obtained using equivalent condoctance and Eq 4.

Modern Theories

The Arrhenius theory explains why solutions of electrolytes conduct electricity, why they exhibit enhanced colligative properties and essentially is satisfactory for solutions of weak electrolytes. Several deficiencies, however, do exist when it is applied to solutions of strong electrolytes. It does not explain the failure of strong electrolytes to follow the law of mass action as applied to ionization; discrepancies exist between the degree of ionization calculated from the van't Hoff factor and the conductivity ratio for strong electrolyte solutions having concentrations greater than about 0.5 *M*.

These deficiencies can be explained hy the following observations

1. In the molten state, strong electrolytes are excellent conductors of electricity. This suggests that these materials are already ionized in the crystalline state. Further support for this is given by X-ray studies of crystals, which indicate that the units comprising the basic lattice structure of strong electrolytes are ions.
2. Arrhenius neglected the fact that ions in solution, being oppositely charged, tend to associate through electrostatic attraction. In solutions of weak electrolytes, the number of ions is not large and it is not surprising that electrostatic attractions do not cause appreciable deviations from theory. In dilute solutions, in which strong electrolytes are assumed to be 100% ionized, the number of ions is large, and interionic attractions become major factors in determining the chemical properties of these solutions. These effects should, and do, become more pronounced as the concentration of electrolyte or the valence of the ions is increased.

It is not surprising, therefore, that the Arrhenius theory of partial ionization involving the law of mass action and neglecting ionic charge does not hold for solutions of strong electrolytes. Neutral molecules of strong electrolytes, if they do exist in solution, must arise from interionic attraction rather than from incomplete ionization.

Activity and Activity Coefficients—Due to increased electrostatic attractions as a solution becomes more concentrated, the concentration of an ion becomes less efficient as a measure of its net effectiveness. A more efficient measure of the physical or chemical effectiveness of an ion is known as its *activity*, which is a measure of an ion's concentration related to its concentration at a universally adopted reference-standard state. The relationship between the activity and the concentration of an ion can be expressed as

$$a = m\gamma \tag{6}$$

where m is the molal concentration, γ is the activity coefficient and a is the activity. The activity also can be expressed in terms of molar concentration, c, as

$$a = fc \tag{7}$$

where f is the activity coefficient on a molar scale. In dilute solutions (below 0.01 *M*) the two activity coefficients are identical, for all practical purposes.

The activity coefficient may be determined in various ways, such as measuring colligative properties, electromotive force, solubility, distribution coefficients, etc. For a strong electrolyte, the mean ionic activity coefficient, $\gamma_\pm$, or $f_\pm$, provides a measure of the deviation of the electrolyte from ideal behavior. The mean ionic activity coefficients on a molal basis for several strong electrolytes are given in Tahle II. It is characteristic of the electrolytes that the coefficients at first decrease with increasing concentration, pass through a minimum and finally increase with increasing concentration of electrolyte.

Ionic Strength—Ionic strength is a measure of the intensity of the electrical field in a solution and may be expressed as

$$\mu = \tfrac{1}{2}\Sigma c_i z_i^2 \tag{8}$$

where z_i is the valence of ion i. The mean ionic activity coefficient is a function of ionic strength as are such diverse phenomena as solubilities of sparingly soluhle substances, rates of ionic reactions, effects of salts on pH of buffers, electrophoresis of proteins, etc.

The greater effectiveness of ions of higher charge type on a specific property, compared with the effectiveness of the same number of singly charged ions, generally coincides with the ionic strength calculated by Eq 8. The variation of ionic strength with the valence (charge) of the ions comprising a strong electrolyte should he noted.

For univalent cations and univalent anions (called uniunivalent or 1-1) electrolytes, the ionic strength is identical with molarity. For bivalent cation and univalent anion (biunivalent or 2-1) electrolytes, or univalent cation and bivalent anion (unibivalent or 1-2) electrolytes, the ionic strength is three times the molarity. For bivalent cation and bivalent anion (bibivalent or 2-2) electrolytes, the ionic strength is four times the molarity. These relationships are evident from the following example.

Example—Calculate the ionic strength of 0.1 *M* solutions of NaCl, Na_2SO_4, $MgCl_2$ and $MgSO_4$, respectively. For

$$\text{NaCl} \quad \mu = \tfrac{1}{2}(0.1 \times 1^2 + 0.1 \times 1^2) = 0.1$$

$$\text{Na}_2\text{SO}_4 \quad \mu = \tfrac{1}{2}(0.2 \times 1^2 + 0.1 \times 2^2) = 0.3$$

$$\text{MgCl}_2 \quad \mu = \tfrac{1}{2}(0.1 \times 2^2 + 0.2 \times 1^2) = 0.3$$

$$\text{MgSO}_4 \quad \mu = \tfrac{1}{2}(0.1 \times 2^2 + 0.1 \times 2^2) = 0.4$$

The ionic strength of a solution containing more than one electrolyte is the sum of the ionic strengths of the individual salts comprising the solution. For example, the ionic strength of a solution containing NaCl, Na_2SO_4, $MgCl_2$ and $MgSO_4$, each at a concentration of 0.1 *M*, is 1.1.

Debye-Hückel Theory—The Dehye-Hückel equations which are applicable only to very dilute solutions (about 0.02 μ), may be extended to somewhat more concentrated solutions (about 0.1 μ) in the simplified form

$$\log f_i = \frac{-0.51\, z_i^2 \sqrt{\mu}}{1 + \sqrt{\mu}} \tag{9}$$

Table III—Values of Some Salting-Out Constants for Various Barbiturates at 25°

Barbiturate	KCl	KBr	NaCl	NaBr
Amobarbital	0.168	0.095	0.212	0.143
Aprobarbital	0.136	0.062	0.184	0.120
Barbital	0.092	0.042	0.136	0.088
Phenobarbital	0.092	0.034	0.132	0.078
Vinbarbital	0.125	0.036	0.143	0.096

The mean ionic activity coefficient for aqueous solutions of electrolytes at 25° can be expressed as

$$\log f_{\pm} = \frac{-0.51\, z_+ z_- \sqrt{\mu}}{1 + \sqrt{\mu}} \qquad (10)$$

in which z_+ is the valence of the cation and z_- is the valence of the anion. When the ionic strength of the solution becomes high (approximately 0.3 to 0.5), these equations become inadequate and a linear term in μ is added. This is illustrated for the mean ionic activity coefficient

$$\log f_{\pm} = \frac{-0.51\, z_+ z_- \sqrt{\mu}}{1 + \sqrt{\mu}} + K_s \mu \qquad (11)$$

in which K_s is a "salting-out" constant chosen empirically for each salt. This equation is valid for solutions with ionic strength up to approximately 1.

Salting-Out Effect—The aqueous solubility of a slightly soluble organic substance generally is affected markedly by the addition of an electrolyte. This effect particularly is noticeable when the electrolyte concentration reaches 0.5 *M* or higher. If the aqueous solution of the organic substance has a dielectric constant lower than that of pure water, its solubility is decreased and the substance is "salted-out." The use of high concentrations of electrolytes, such as ammonium sulfate or sodium sulfate, for the separation of proteins by differential precipitation is perhaps the most striking example of this effect. The aqueous solutions of a few substances such as hydrocyanic acid, glycine and cystine have a higher dielectric constant than that of pure water, and these substances are "salted-in." These phenomena can be expressed empirically as

$$\log S = \log S_0 \pm K_s m \qquad (12)$$

in which S_0 represents the solubility of the organic substance in pure water and S is the solubility in the electrolyte solution. The slope of the straight line obtained by plotting log S versus m is positive for "salting-in" and negative for salting-out. In terms of ionic strength this equation becomes

$$\log S = \log S_0 \pm K_s' \mu \qquad (13)$$

where $K_s' = K_s$ for univalent salts, $K_s' = K_s/3$ for unibivalent salts and $K_s' = K_s/4$ for bivalent salts. The salting-out constant depends on the temperature as well as the nature of both the organic substance and the electrolyte. The effect of the electrolyte and the organic substance can be seen in Table III. In all instances, if the anion is constant, the sodium cation has a greater salting-out effect than the potassium cation, probably due to the higher charge density of the former. Although the reasoning is less clear, it appears that for a constant cation, chloride anion has a greater effect than bromide anion upon the salting-out phenomenon.

Acids and Bases

Arrhenius defined an acid as a substance that yields hydrogen ions in aqueous solution and a base as a substance that yields hydroxyl ions in aqueous solution. Except for the fact that hydrogen ions neutralize hydroxyl ions to form water, no complementary relationship between acids and bases (such as that between oxidants and reductants, for example) is evident in Arrhenius' definitions for these substances; rather, their oppositeness of character is emphasized. Moreover, no account is taken of the behavior of acids and bases in nonaqueous solvents. Also, while acidity is associated with so elementary a particle as the proton (hydrogen ion), basicity is attributed to so relatively complex an association of atoms as the hydroxyl ion. It would seem that a simpler concept of a base could be devised.

Proton Concept—In pondering the objections to Arrhenius' definitions, Brønsted and Bjerrum in Denmark and Lowry in England developed, and in 1923 announced, a more satisfactory, and more general, theory of acids and bases. According to this theory an acid is a substance capable of yielding a proton (hydrogen ion), while a base is a substance capable of accepting a proton. This complementary relationship may be expressed by the general equation

$$\underset{\text{acid}}{A} \rightleftharpoons H^+ + \underset{\text{base}}{B}$$

The pair of substances thus related through mutual ability to gain or lose a proton is called a *conjugate acid–base pair*. Specific examples of such pairs are

Acid		Base
HCl	$\rightleftharpoons$	$H^+ + Cl^-$
CH_3COOH	$\rightleftharpoons$	$H^+ + CH_3COO^-$
NH_4^+	$\rightleftharpoons$	$H^+ + NH_3$
HCO_3^-	$\rightleftharpoons$	$H^+ + CO_3^{2-}$
$H_2PO_4^-$	$\rightleftharpoons$	$H^+ + HPO_4^{2-}$
H_2O	$\rightleftharpoons$	$H^+ + OH^-$
H_3O^+	$\rightleftharpoons$	$H^+ + H_2O$
$Al(H_2O)_6^{3+}$	$\rightleftharpoons$	$H^+ + Al(H_2O)_5OH^{2+}$

It is apparent that not only molecules but also cations and anions may function as acids or bases.

The complementary nature of the acid–base pairs listed is reminiscent of the complementary relationship of pairs of oxidants and reductants where, however, the ability to gain or lose one or more electrons—rather than protons—is the distinguishing characteristic.

Oxidant		Reductant
$Fe^{3+} + e^-$	$\rightleftharpoons$	Fe^{2+}
$Na^+ + e^-$	$\rightleftharpoons$	Na
$\frac{1}{2}I_2 + e^-$	$\rightleftharpoons$	I^-

However, these examples of acid–base pairs and oxidant–reductant pairs represent reactions that are possible in principle only. Ordinarily acids will not release free protons any more than reductants will release free electrons. That is, protons and electrons, respectively, can be *transferred* only from one substance (an ion, atom or molecule) to another. Thus, it is a fundamental fact of chemistry that oxidation of one substance will occur only if reduction of another substance occurs simultaneously. Stated in another way, electrons will be released from the reductant (oxidation) only if an oxidant capable of accepting electrons (reduction) is present. For this reason oxidation–reduction reactions must involve two conjugate oxidant–reductant pairs of substances

$$\text{oxidant}_1 + \text{reductant}_2 \rightleftharpoons \text{reductant}_1 + \text{oxidant}_2$$

where Subscript 1 represents one conjugate oxidant–reductant pair and Subscript 2 represents the other.

Similarly, an acid will not release a proton unless a base capable of accepting it is present simultaneously. This means that any actual manifestation of acid–base behavior must involve interaction between two sets of conjugate acid–base pairs, represented as

$$\underset{\text{acid}_1}{A_1} + \underset{\text{base}_2}{B_2} \rightleftharpoons \underset{\text{base}_1}{B_1} + \underset{\text{acid}_2}{A_2}$$

In such a reaction, which is called *protolysis* or a *protolytic reaction*, A_1 and B_1 constitute one conjugate acid–base pair and A_2 and B_2 the other; the proton given up by A_1 (which thereby becomes B_1) is transferred to B_2 (which becomes A_2).

When an acid, such as hydrochloric, is dissolved in water, a protolytic reaction occurs.

$$\underset{\text{acid}_1}{HCl} + \underset{\text{base}_2}{H_2O} \rightleftharpoons \underset{\text{base}_1}{Cl^-} + \underset{\text{acid}_2}{H_3O^+}$$

The ionic species H_3O^+, called *hydronium* or *oxonium* ion, always is formed when an acid is dissolved in water. Often, for purposes of convenience, this is written simply as H^+ and is called hydrogen ion, although the "bare" ion practically is nonexistent in solution.

When a base, eg, ammonia, is dissolved in water the reaction of protolysis is

$$\underset{\text{base}_1}{NH_3} + \underset{\text{acid}_2}{H_2O} \rightleftharpoons \underset{\text{acid}_1}{NH_4^+} + \underset{\text{base}_2}{OH^-}$$

The proton theory of acid–base function makes the concept of hydrolysis superfluous. When, for example, sodium acetate is dissolved in water, this acid–base interaction occurs

$$\underset{\text{base}_1}{CH_3COO^-} + \underset{\text{acid}_2}{H_2O} \rightleftharpoons \underset{\text{acid}_1}{CH_3COOH} + \underset{\text{base}_2}{OH^-}$$

In an aqueous solution of ammonium chloride the reaction is

$$\underset{\text{acid}_1}{NH_4^+} + \underset{\text{base}_2}{H_2O} \rightleftharpoons \underset{\text{base}_1}{NH_3} + \underset{\text{acid}_2}{H_3O^+}$$

Transfer of protons (protolysis) is not limited to dissimilar conjugate acid–base pairs. In the preceding examples H_2O sometimes behaves as an acid and at other times as a base. Such an amphoteric substance is called, in Brønsted's terminology, an *amphiprotic substance*.

Electron-Pair Concept—While the proton concept of acids and bases provides a more general definition for these substances, it does not indicate the basic reason for proton transfer, nor does it explain how such substances as sulfur trioxide, boron trichloride, stannic chloride or carbon dioxide—none of which is capable of donating a proton—can behave as acids. Both deficiencies of the proton theory are avoided in the more inclusive definition of acids and bases proposed by Lewis in 1923. In 1916 he proposed that sharing of a pair of electrons by two atoms established a bond (covalent) between the atoms; therefore, an acid is a substance capable of sharing a pair of electrons made available by another substance called a base, thereby forming a coordinate covalent bond. The base is the substance that donates a share in its electron pair to the acid. The following equation illustrates how Lewis' definitions explain the transfer of a proton (hydrogen ion) to ammonia to form ammonium ion.

$$H^+ + :\ddot{N}:H \rightarrow \left[H:\ddot{N}:H \right]^+$$

The reaction of boron trichloride, which according to the Lewis theory is an acid, with ammonia is similar, for the boron lacks an electron pair if it is to attain a stable octet configuration, while ammonia has a pair of electrons which may be shared, thus

$$Cl:\ddot{B}:Cl + :\ddot{N}:H \rightarrow Cl:\ddot{B}:\ddot{N}:H$$

Leveling Effect of a Solvent—When the strong acids such as $HClO_4$, H_2SO_4, HCl or HNO_3 are dissolved in water the solutions—if they are of identical normality and are not too concentrated—all have about the same hydrogen-ion concentration, indicating the acids to be of about the same strength. The reason for this is that each one of the acids undergoes practically complete protolysis in water.

$$\underset{\text{acid}_1}{HCl} + \underset{\text{base}_2}{H_2O} \rightarrow \underset{\text{base}_1}{Cl^-} + \underset{\text{acid}_2}{H_3O^+}$$

This phenomenon, called the leveling effect of water, occurs whenever the added acid is stronger than the hydronium ion. Such a reaction manifests the tendency of proton-transfer reactions to proceed spontaneously in the direction of forming a weaker acid or weaker base.

Since the strongest acid that can exist in an amphiprotic solvent is the conjugate acid form of the solvent, any stronger acid will undergo protolysis to the weaker solvent acid. Since $HClO_4$, H_2SO_4, HCl or HNO_3 are all stronger acids than the hydronium ion, they are converted in water to the hydronium ion.

When the strong bases sodium hydride, sodium amide or sodium ethoxide are dissolved in water, each reacts with water to form sodium hydroxide. These reactions illustrate the leveling effect of water on bases. Since the hydroxide ion is the strongest base which can exist in water, any base stronger than the hydroxide ion undergoes protolysis to hydroxide.

Intrinsic differences in the acidity of acids become evident if they are dissolved in a relatively poor proton acceptor such as anhydrous acetic acid. Perchloric acid ($HClO_4$), a strong acid, undergoes practically complete reaction with acetic acid

$$\underset{\substack{\text{acid}_1\\(\text{strong})}}{HClO_4} + \underset{\substack{\text{base}_2\\(\text{strong})}}{CH_3COOH} \rightarrow \underset{\substack{\text{base}_1\\(\text{weak})}}{ClO_4^-} + \underset{\substack{\text{acid}_2\\(\text{weak})}}{CH_3COOH_2^+}$$

but sulfuric acid and hydrochloric acid behave as weak acids. It is because perchloric acid is a very strong acid when dissolved in glacial acetic acid that it has found many important applications in analytical chemistry as a titrant for a variety of substances which behave as bases in acetic acid. Because of its ability to differentiate the acidity of various acids, it is called a *differentiating solvent for acids*; this property results from its relatively weak proton-acceptor tendency. A solvent that differentiates basicity of different hases must bave a weak proton-donor tendency; it is called a *differentiating solvent for bases*. Typical of solvents in this category is liquid ammonia. Solvents that have both weak proton-donor and proton-acceptor tendencies are called *aprotic solvents* and may serve as differentiating solvents for both acids and bases; they have little if any action on solutes and serve mainly as inert dispersion media for the solutes. Useful aprotic solvents are benzene, toluene or hexane.

Ionization of Acids and Bases—Acids and bases commonly are classified as strong or weak acids and strong or

weak bases depending on whether they are ionized extensively or slightly in aqueous solutions. If, for example, 1 *N* aqueous solutions of hydrochloric acid and acetic acid are compared, it is found that the former is a better conductor of electricity, reacts much more readily with metals, catalyzes certain reactions more efficiently and possesses a more acid taste than the latter. Both solutions, however, will neutralize identical amounts of alkali. A similar comparison of 1 *N* solutions of sodium hydroxide and ammonia reveals the former to be more "active" than the latter, although both solutions will neutralize identical quantities of acid.

The differences in the properties of the two acids is attributed to differences in the concentration of hydrogen (more accurately hydronium) ion, the hydrochloric acid being ionized to a greater extent and, therefore, containing a higher concentration of hydrogen ion than acetic acid. Similarly, most of the differences between the sodium hydroxide and ammonia solutions are attributed to the higher hydroxyl-ion concentration in the former.

The ionization of incompletely ionized acids may be considered a reversible reaction of the type

$$HA \rightleftharpoons H^+ + A^-$$

where HA is the molecular acid and A^- is its anion. An equilibrium expression based on the law of mass action may be applied to the reaction

$$K_a = \frac{[H^+][A^-]}{[HA]} \tag{14}$$

where K_a is the ionization or dissociation constant, and the brackets signify concentration. For any given acid in any specified solvent and at any constant temperature, K_a remains relatively constant as the concentration of acid is varied, provided the acid is weakly ionized. With increasingly stronger acids, however, progressively larger deviations occur.

Although the strength of an acid commonly is measured in terms of the ionization or dissociation constant defined in Eq 14, the process of ionization probably is never as simple as shown above. A proton simply will not detach itself from one molecule unless it is accepted simultaneously by another molecule. When an acid is dissolved in water, the latter acts as a base, accepting a proton (Brønsted's definition of a base) by donating a share in a pair of electrons (Lewis' definition of a base). This reaction may be written

$$\underset{\text{acid}_1}{HA} + \underset{\text{base}_2}{H_2O} \rightleftharpoons \underset{\text{base}_1}{A^-} + \underset{\text{acid}_2}{H_3O^+}$$

Application of the law of mass action to this reaction gives

$$K = \frac{[H_3O^+][A^-]}{[HA][H_2O]} \tag{15}$$

since $[H_2O]$ is a constant this equation may be written

$$K_a = \frac{[H_3O^+][A^-]}{[HA]} \tag{16}$$

This equation is identical with Eq 14 because $[H_3O^+]$ is numerically equal to $[H^+]$.

Acids which are capable of donating more than one proton are termed *polyprotic*. The ionization of a polyprotic acid occurs in stages and can be illustrated by considering the equilibria involved in the ionization of phosphoric acid

$$H_3PO_4 + H_2O \rightleftharpoons H_2PO_4^- + H_3O^+$$

$$H_2PO_4^- + H_2O \rightleftharpoons HPO_4^{2-} + H_3O^+$$

$$HPO_4^{2-} + H_2O \rightleftharpoons PO_4^{3-} + H_3O^+$$

Application of the law of mass action to this series of reactions gives

$$K_1 = \frac{[H_2PO_4^-][H_3O^+]}{[H_3PO_4]} \tag{17}$$

$$K_2 = \frac{[HPO_4^{2-}][H_3O^+]}{[H_2PO_4^-]} \tag{18}$$

$$K_3 = \frac{[PO_4^{3-}][H_3O^+]}{[HPO_4^{2-}]} \tag{19}$$

If the three expressions for the ionization constants are multiplied together, an overall ionization, K, can be obtained

$$K = K_1K_2K_3 = \frac{[PO_4^{3-}][H_3O^+]^3}{[H_3PO_4]} \tag{20}$$

Each of the successive ionizations is suppressed by the hydronium ion formed from preceding stages according to Le Chatelier's principle. The successive dissociation constants always decrease in value, since successive protons must be removed from species that are always more negatively charged. This can be seen from the data in Table IV, in which K_1 for phosphoric acid is approximately 100,000 times greater than K_2, which is in turn approximately 100,000 times greater than K_3. Although successive dissociation constants are always smaller, the difference is not always as great as it is for phosphoric acid. Tartaric acid, for example, has $K_1 = 9.12 \times 10^{-4}$ and $K_2 = 4.27 \times 10^{-5}$.

Ionization of a base can be illustrated by using the specific substance NH_3 for an example. According to Brønsted and Lewis, when the base NH_3 is dissolved in water, the latter acts as an acid, donating a proton to NH_3, which accepts it by offering a share in a pair of electrons on the nitrogen atom. This reaction is written

$$\underset{\text{base}}{NH_3} + \underset{\text{acid}}{H_2O} \rightleftharpoons NH_4^+ + OH^-$$

The equilibrium expression for this reaction is

$$K = \frac{[NH_4^+][OH^-]}{[NH_3][H_2O]} \tag{21}$$

with $[H_2O]$ constant this expression may be written

$$K_b = \frac{[NH_4^+][OH^-]}{[NH_3]} \tag{22}$$

Ionization of Water—Although it is a poor conductor of electricity, pure water does ionize through a process known as *autoprotolysis*, in the following manner

$$2H_2O \rightleftharpoons H_3O^+ + OH^-$$

Application of the law of mass action to this reaction gives

$$K = \frac{[H_3O^+][OH^-]}{[H_2O]^2} \tag{23}$$

where K is the equilibrium constant for the reaction. Since the concentration of H_2O (molecular water) is very much greater than either the hydronium-ion or hydroxyl-ion concentrations, it can be considered to be constant and can be combined with K to give a new constant, K_w, known as the *ion product* of water, and Eq 23 becomes

$$K_w = [H_3O^+][OH^-] \tag{24}$$

The numerical value of K_w varies with temperature; at 25° it is approximately equal to 1×10^{-14}.

Since the autoprotolysis of pure water yields one hydronium ion for each hydroxyl ion produced, $[H_3O^+]$ must be equal to $[OH^-]$. At 25° each has a value of 1×10^{-7} moles/

Table IV—Dissociation Constants in Water at 25°

Substance		K
Weak Acids		
Acetic		1.75×10^{-5}
Acetylsalicylic		3.27×10^{-4}
Barbital		1.23×10^{-8}
Barbituric		1.05×10^{-4}
Benzoic		6.30×10^{-5}
Benzyl penicillin		1.74×10^{-3}
Boric	K_1	5.8×10^{-10}
Caffeine		1×10^{-14}
Carbonic	K_1	4.31×10^{-7}
	K_2	4.7×10^{-11}
Citric ($1H_2O$)	K_1	7.0×10^{-4}
	K_2	1.8×10^{-5}
	K_3	4.0×10^{-7}
Dichloroacetic		5×10^{-2}
Ethylenediaminetetra-	K_1	1×10^{-2}
acetic acid (EDTA)	K_2	2.14×10^{-3}
	K_3	6.92×10^{-7}
	K_4	5.5×10^{-11}
Formic		1.77×10^{-4}
Glycerophosphoric	K_1	3.4×10^{-2}
	K_2	6.4×10^{-7}
Glycine	K_1	4.5×10^{-3}
	K_2	1.7×10^{-10}
Lactic		1.39×10^{-4}
Mandelic		4.29×10^{-4}
Monochloroacetic		1.4×10^{-3}
Oxalic ($2H_2O$)	K_1	5.5×10^{-2}
	K_2	5.3×10^{-5}
Phenobarbital		3.9×10^{-8}
Phenol		1×10^{-10}
Phosphoric	K_1	7.5×10^{-3}
	K_2	6.2×10^{-8}
	K_3	2.1×10^{-13}
Picric		4.2×10^{-1}
Propionic		1.34×10^{-5}
Saccharin		2.5×10^{-2}
Salicylic		1.06×10^{-3}
Succinic	K_1	6.4×10^{-5}
	K_2	2.3×10^{-6}
Sulfadiazine		3.3×10^{-7}
Sulfamerazine		8.7×10^{-8}
Sulfapyridine		3.6×10^{-9}
Sulfathiazole		7.6×10^{-8}
Tartaric	K_1	9.6×10^{-4}
	K_2	4.4×10^{-5}
Trichloroacetic		1.3×10^{-1}
Weak Bases		
Acetanilide		4.1×10^{-14} (40°)
Ammonia		1.74×10^{-5}
Apomorphine		1.0×10^{-7}
Atropine		4.5×10^{-5}
Benzocaine		6.0×10^{-12}
Caffeine		4.1×10^{-14} (40°)
Cocaine		2.6×10^{-6}
Codeine		9×10^{-7}
Ephedrine		2.3×10^{-5}
Morphine		7.4×10^{-7}
Papaverine		8×10^{-9}
Physostigmine	K_1	7.6×10^{-7}
	K_2	5.7×10^{-13}
Pilocarpine	K_1	7×10^{-8}
	K_2	2×10^{-13}
Procaine		7×10^{-6}
Pyridine		1.4×10^{-9}
Quinine	K_1	1.0×10^{-6}
	K_2	1.3×10^{-10}
Reserpine		4×10^{-8}
Strychnine	K_1	1×10^{-6}
	K_2	2×10^{-12}
Theobromine		4.8×10^{-14} (40°)
Thiourea		1.1×10^{-15}
Urea		1.5×10^{-14}

liter ($1 \times 10^{-7} \times 1 \times 10^{-7} = K_w = 1 \times 10^{-14}$). A solution in which $[H_3O^+]$ is equal to $[OH^-]$ is termed a *neutral* solution.

If an acid is added to water, the hydronium-ion concentration will be increased and the equilibrium between hydronium and hydroxyl ions will be disturbed *momentarily*. To restore equilibrium, some of the hydroxyl ions, originally present in the water, will combine with a *part* of the added hydronium ions to form nonionized water molecules, until the product of the concentrations of the two ions has been reduced to 10^{-14}. When equilibrium again is restored, the concentrations of the two ions no longer will be equal. If, for example, the hydronium-ion concentration is 1×10^{-3} N when equilibrium is established, the concentration of hydroxyl ion will be 1×10^{-11} (the product of the two concentrations being equal to 10^{-14}). Since $[H_3O^+]$ is much greater than $[OH^-]$, the solution is said to be *acid* or *acidic*.

In a similar manner, the addition of an alkali to pure water momentarily disturbs the equilibrium between hydronium and hydroxyl ions. To restore equilibrium, some of the hydronium ions originally present in the water will combine with part of the added hydroxyl ions to form nonionized water molecules. The process continues until the product of the hydronium and hydroxyl ion concentrations again is equal to 10^{-14}. Assuming that the final hydroxyl-ion concentration is 1×10^{-4} N, the concentration of hydronium ion in the solution will be 1×10^{-10}. Since $[OH^-]$ is much greater than $[H_3O^+]$, the solution is said to be *basic* or *alkaline*.

Relationship of K_a and K_b—A particularly interesting and useful relationship between the strength of an acid and its conjugate base, or a base and its conjugate acid, exists. For illustration, consider the strength of the base NH_3 and its conjugate acid NH_4^+, in water. The behavior of NH_3 as a base is expressed by

$$NH_3 + H_2O \rightleftharpoons NH_4^+ + OH^-$$

for which the equilibrium, as formulated earlier is

$$K_b = \frac{[NH_4^+][OH^-]}{[NH_3]} \tag{25}$$

The behavior of NH_4^+ as an acid is represented by

$$NH_4^+ + H_2O \rightleftharpoons NH_3 + H_3O^+$$

the equilibrium constant for which is

$$K_a = \frac{[NH_3][H_3O^+]}{[NH_4^+]} \tag{26}$$

Multiplying Eqs 25 and 26

$$K_aK_b = \frac{[NH_3][H_3O^+][NH_4^+][OH^-]}{[NH_4^+][NH_3]} \tag{27}$$

It is obvious that the product

$$K_w = K_a.K_b \tag{28}$$

where K_w is the ion product of water as defined in Eq 24.

The utility of this relationship, which is a general one for any conjugate acid–base pair, is evident from the following deductions: (1) the strength of an acid may be expressed in terms either of the K_a or the K_b of its conjugate base, or *vice versa;* (2) the K_a of an acid may be calculated if the K_b of its conjugate base is known, or *vice versa* and (3) the stronger an acid is, the weaker its conjugate base, or *vice versa*.

Bases which are capable of interacting with more than one proton are termed *polyacidic*, and can be illustrated by

$$PO_4^{3-} + H_2O \rightleftharpoons HPO_4^{2-} + OH^-$$

$$HPO_4^{2-} + H_2O \rightleftharpoons H_2PO_4^- + OH^-$$

$$H_2PO_4^- + H_2O \rightleftharpoons H_3PO_4 + OH^-$$

Applying the law of mass action to this series of reactions, and using the concepts outlined in Eqs 25 to 28, it becomes obvious that the relationship between the various K_a and K_b values for phosphoric acid are

$$K_w = K_{a1} \times K_{b3} = K_{a2} \times K_{b2} = K_{a3} \times K_{b1} \qquad (29)$$

where K_{a1}, K_{a2} and K_{a3} refer to the equilibria given by Eqs 17, 18 and 19, respectively; K_{b1}, K_{b2} and K_{b3} refer to the reaction of PO_4^{3-}, HPO_4^{2-} and $H_2PO_4^-$, respectively, with water.

Electronegativity and Dissociation Constants—Table IV gives the dissociation constants of several weak acids and weak bases, in water, at 25°. As pointed out previously, strong acids and strong bases do not obey the law of mass action, so that dissociation constants cannot be formulated for these strong electrolytes.

From an inspection of this table it is evident that great variations occur in the strength of weak acids and weak bases. The effect of various substituents on the strength of acids and bases depends on the electronegativity of the substituent atom or radical. For example, the substitution of one chlorine atom into the molecule of acetic acid increases the degree of ionization of the acid. Substitution of two chlorine atoms further increases the degree of ionization, and introduction of three chlorine atoms produces a still stronger acid. Acetic acid ionizes primarily because the oxygen atom adjacent to the hydrogen atom of the carboxyl group has a stronger affinity for electrons than has the hydrogen atom. Thus, when acetic acid is dissolved in water the polar molecules of the latter have a stronger affinity for the hydrogen of acetic acid than the latter. The acetic acid ionizes as a consequence of this difference in affinities. When an atom of chlorine is introduced into the acetic acid molecule, forming $ClCH_2COOH$, the electrons in the molecule are attracted very strongly to the chlorine because of its relatively high electronegativity, the bond between the hydrogen and the oxygen in the carboxyl group thereby weakened and the degree of ionization increased. Introduction of two, or three chlorine atoms weakens the bond further and increases the strength of the acid. On the other hand, substitution of chlorine into the molecule of ammonia reduces the strength of the base because of its decreased affinity for the hydrogen ion.

Ionic Strength and Dissociation Constants—Most solutions of pharmaceutical interest are in a concentration range such that the ionic strength of the solution may have a marked effect on ionic equilibria and observed dissociation constants. One method of correcting dissociation constants for solutions with an ionic strength up to about 0.3 is to calculate an apparent dissociation constant, pK_a' as

$$pK_a' = pK_a + \frac{0.51\,(2Z-1)\sqrt{\mu}}{1+\sqrt{\mu}} \qquad (30)$$

in which pK_a is the tabulated thermodynamic dissociation constant, Z is the charge on the acid and μ is the ionic strength.

Example—Calculate pK_2' for succinic acid at an ionic strength of 0.1. Assume that pK_2 is 5.63. The charge on the acid species is −1.

$$pK_2' = 5.63 + \frac{0.51\,(-2-1)\sqrt{0.1}}{1+\sqrt{0.1}}$$

$$= 5.63 - 0.37 = 5.26$$

Determination of Dissociation Constants—Although the dissociation constant of a weak acid or base can be obtained in a wide variety of ways including conductivity measurements, or ultraviolet or visible absorption spectrometry, the most widely used method is potentiometric pH measurement (see *Potentiometry*, page 244). The simplest method involving potentiometric pH measurement is based on the measurement of the hydronium-ion concentration of a solution containing equimolar concentrations of the acid and a strong-base salt of the acid. The principle of this method is evident from an inspection of Eq 16; when equimolar concentrations of HA (the acid) and A^- (the salt) are present, the dissociation constant, K_a, numerically is equal to the hydronium-ion concentration (also, the pK_a of the acid is equal to the pH of the solution). Although this method is simple and rapid, the dissociation constant obtained is not sufficiently accurate for many purposes.

In order to obtain the dissociation constant of a weak acid with a high degree of accuracy and precision, a dilute solution of the acid (about 10^{-3} to 10^{-4} M) is titrated with a strong base, and the pH of the solution taken after each addition of base. The resulting data can be handled in a wide variety of ways, perhaps the best of which is the method proposed by Benet and Goyan.[1] The proton balance equation for a weak acid, HA, being titrated with a strong base such as KOH, would be

$$[K^+] + [H_3O^+] = [OH^-] + [A^-] \qquad (31)$$

in which $[K^+]$ is the concentration of the base added. Equation 31 can be rearranged to give

$$Z = [A^-] = [K^+] + [H_3O^+] - [OH^-] \qquad (32)$$

When a weak monoprotic acid is added to water, it can exist in the unionized form, HA, and in the ionized form, A^-. After equilibrium is established, the sum of the concentrations of both species must be equal to C_a, the stoichiometric (added) concentration of acid or

$$C_a = [HA] + [A^-] = [HA] + Z \qquad (33)$$

The term, [HA], can be replaced using Eq 16 to give

$$C_a = \frac{[H_3O^+]Z}{K_a} + Z \qquad (34)$$

which can be rearranged to

$$Z = C_a - \frac{Z[H_3O^+]}{K_a} \qquad (35)$$

According to Eq 35, if Z, which is obtained from the experimental data using Eq 32, is plotted versus the terms $Z[H_3O^+]$, a straight line results with a slope equal to $1/K_a$, and an intercept equal to C_a. In addition to obtaining an accurate estimate for the dissociation constant, the stoichiometric concentration of the substance being titrated is obtained also. This is of importance when the substance being titrated cannot he purified, or has an unknown degree of solvation. Similar equations can be developed for ohtaining the dissociation constant for a weak base.[1]

The dissociation constants for diprotic acids can be ohtained by defining P as the average number of protons dissociated per mole of acid or

$$P = Z/C_a \qquad (36)$$

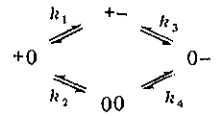
and

$$\frac{[H_3O^+]^2P}{(2-P)} = K_1K_2 + \frac{K_1[H_3O^+]\,(1-P)}{(2-P)} \qquad (37)$$

A plot of Eq 37 should yield a straight line with a slope equal to K_1 and an intercept of K_2; dividing the intercept by the slope yields K_2.

Micro Dissociation Constants—The dissociation constants for polyprotic acids, as determined hy potentiometric titration, are generally known as macro, or titration con-

stants. Since it is known that carboxyl groups are stronger acids than protonated amino groups, there is no difficulty in assigning K_1 and K_2, as determined by Eq 37 to the carboxyl and amino groups, respectively, of a substance such as glycine hydrochloride. In other chemicals or drugs such as phenylpropanolamine, in which the two acidic groups are the phenolic and the protonated amino group, the assignment of dissociation constants is more difficult. This is because, in general, both groups have dissociation constants of equal magnitude. Thus, there will be two ways of losing the first proton and two ways of losing the second, resulting in four possible species in solution. This can be illustrated using the convention of assigning a + to a positively charged group, a 0 to an uncharged group and a − to a negatively charged group. Thus, +0 would represent the fully protonated phenylpropanolamine, +− the dipolar ion, 00 the uncharged molecule and 0−, the anion. The total ionization scheme, therefore, can be written



The micro constants are related to the macro constants as

$$K_1 = k_1 + k_2 \tag{38}$$

$$K_1K_2 = k_1k_3 = k_2k_4 \tag{39}$$

It can be seen from Eq 38 that unless k_1 or k_2 is very much smaller than the other, the observed macro constant is a composite of the two and cannot he assigned to one or the other acidic group in a nonambiguous way.

Methods for determining k_1 are given by Riegelman[2] and Niebergall *et al.*[3] Once k_1, K_1 and K_2 have been determined, all of the other micro constants can be obtained from Eqs 38 and 39.

pH

The numerical values of hydronium-ion concentration may vary enormously; for a normal solution of a strong acid the value is nearly 1, while for a normal solution of a strong base it is approximately 1×10^{-14}; ie, a variation of 100,000,000,000,000 between these two limits. Because of the inconvenience of dealing with such large numbers, Sørenson, in 1909, proposed that hydronium-ion concentration be expressed in terms of the logarithm (log) of its reciprocal. To this value he assigned the symbol pH. Mathematically it is written

$$\text{pH} = \log \frac{1}{[H_3O^+]} \tag{40}$$

and since the logarithm of 1 is zero, the equation also may be written

$$\text{pH} = -\log[H_3O^+] \tag{41}$$

from which it is evident that pH also may be defined as the negative logarithm of tbe hydronium-ion concentration. In general, this type of notation is used to indicate the negative logarithm of the term that is preceded by the "p," which gives rise to the following

$$\text{pOH} = -\log [OH^-] \tag{42}$$

$$\text{p}K = -\log K \tag{43}$$

Thus, taking logarithms of Eqs 26 and 28 gives

$$\text{p}K_a + \text{p}K_b = \text{p}K_w \tag{44}$$

$$\text{pH} + \text{pOH} = \text{p}K_w \tag{45}$$

The relationship of pH to hydronium-ion and hydroxyl-ion concentrations may be seen in Table V.

The following examples illustrate the conversion from exponential to "p" notation

1. Calculate the pH corresponding to a hydronium-ion concentration of 1×10^{-4} g-ion/L.

Solution:

$$\text{pH} = \log \frac{1}{1 \times 10^{-4}}$$

$$= \log 10{,}000 \text{ or } \log (1 \times 10^{+4})$$

$$\log (1 \times 10^{+4}) = +4$$

$$\text{pH} = 4$$

Table V—Hydronium-Ion and Hydroxyl-Ion Concentrations

	pH	Normality in Terms of Hydronium Ion	Normality in Terms of Hydroxyl Ion
	0	1	10^{-14}
	1	10^{-1}	10^{-13}
Increasing acidity	2	10^{-2}	10^{-12}
	3	10^{-3}	10^{-11}
	4	10^{-4}	10^{-10}
	5	10^{-5}	10^{-9}
	6	10^{-6}	10^{-8}
Neutral point	7	10^{-7}	10^{-7}
	8	10^{-8}	10^{-6}
	9	10^{-9}	10^{-5}
Increasing alkalinity	10	10^{-10}	10^{-4}
	11	10^{-11}	10^{-3}
	12	10^{-12}	10^{-2}
	13	10^{-13}	10^{-1}
	14	10^{-14}	1

2. Calculate the pH corresponding to a hydronium ion-concentration of 0.000036 *N* (or g-ion/L) (*Note*—This more frequently is written as a number multiplied by a power of 10, thus, 3.6×10^{-5} for 0.000036.)

Solution

$$\text{pH} = \log \frac{1}{3.6 \times 10^{-5}}$$

$$= \log 28{,}000 \text{ or } \log (2.8 \times 10^{+4})$$

$$\log (2.8 \times 10^{+4}) = \log 2.8 + \log 10^{+4}$$

$$\log 2.8 = +0.44$$

$$\log 10^{+4} = \underline{+4.00}$$

$$\text{pH} = 4.44$$

This problem may also be solved as follows

$$\text{pH} = -\log (3.6 \times 10^{-5})$$

$$\log 3.6 = +0.56$$

$$\log 10^{-5} = \underline{-5.00}$$

$$= -4.44 = \log (3.6 \times 10^{-5})$$

$$\text{pH} = -(-4.44) = +4.44 = 4.44$$

The following examples illustrate the conversion of "p" notation to exponential notation:

1. Calculate the hydronium-ion concentration corresponding to a pH of 4.44.

Solution

$$\text{pH} = \log \frac{1}{[H_3O^+]}$$

$$4.44 = \log \frac{1}{[H_3O^+]}$$

$$\frac{1}{[H_3O^+]} = \text{antilog of } 4.44 = 28{,}000 \text{ (rounded off)}$$

$$[H_3O^+] = \frac{1}{28{,}000} = 0.000036 \text{ or } 3.6 \times 10^{-5}$$

This calculation also may be made as

$$+4.44 = -\log [H_3O^+]$$

$$\text{or } -4.44 = +\log [H_3O^+]$$

In finding the antilog of −4.44 it should be kept in mind that the *mantissa* (the number to the right of the decimal point) of a log to the base 10 (the common or Briggsian logarithm base) is *always positive* but that the *characteristic* (the number to the left of the decimal point) may be *positive* or *negative*. As the entire log −4.44 is negative, it is obvious that one cannot look up the antilog of −0.44. However, the number −4.44 also may be written (−5.00 + 0.56) or, as more often written, $\bar{5}.56$, where the bar across the characteristic indicates that it alone is negative, while the rest of the number is positive. Looking up the antilog of 0.56 it is found to be 3.6 and, as the antilog of −5.00 is 10^{-5}, it follows that the hydronium-ion concentration must be 3.6×10^{-5} moles/L.

2. Calculate the hydronium-ion concentration corresponding to a pH of 10.17.

Solution

$$10.17 = -\log [H_3O^+]$$

$$-10.17 = \log [H_3O^+]$$

$$-10.17 = (-11.00 + 0.83) = \bar{1}\bar{1}.83$$

The antilog of 0.83 = 6.8
The antilog of $-11.00 = 10^{-11}$
The hydronium-ion concentration is therefore 6.8×10^{-11} moles/L.

In the section on *Ionization of Water* it was shown that the hydronium-ion concentration of pure water, at 25°, is 1×10^{-7} *N*, corresponding to a pH of 7.* This figure, therefore, is designated as the neutral point and all values below a pH of 7 represent acidity; the smaller the number, the greater the acidity. Values above 7 represent alkalinity; the larger the number, the greater the alkalinity. The pH scale usually runs from 0 to 14, but mathematically there is no reason why negative numbers or numbers above 14 should not be used. In practice, however, such values are never encountered because solutions which might be expected to have such values are too concentrated to be ionized extensively or the interionic attraction is so great as to materially reduce ionic activity.

It should be emphasized strongly that the generalizations stated concerning neutrality, acidity and alkalinity hold exactly only when (1) the solvent is water, (2) the temperature is 25° and (3) there are no other factors to cause deviation from the simply formulated equilibria underlying the definition of pH given in the preceding discussion.

Species Concentration

When a weak acid, H_nA, is added to water, $n + 1$ species, including the unionized acid, can exist. After equilibrium is established, the sum of the concentrations of all species must be equal to C_a, the stoichiometric (added) concentration of acid. Thus, for a triprotic acid H_3A

$$C_a = [H_3A] + [H_2A^-] + [HA^{2-}] + [A^{3-}] \tag{46}$$

In addition, the concentrations of all acidic and basic species in solution vary with pH, and can be represented solely in terms of equilibrium constants and the hydronium-ion concentration. These relationships may be expressed as

$$[H_nA] = [H_3O^+]^n C_a/D \tag{47}$$

$$[H_{n-j}A^{-j}] = [H_3O^+]^{n-j} K_1 \ldots K_j C_a/D \tag{48}$$

in which n represents the total number of dissociable hydrogens in the parent acid, j is the number of protons dissociated, C_a is the stoichiometric concentration of acid and K represents the acid dissociation constants. The term D is a power series in $[H_3O^+]$ and K, starting with $[H_3O^+]$ raised to the nth power. The last term is the product of all the dissociation constants. The intermediate terms can be generated from the last term by substituting $[H_3O^+]$ for K_n to obtain the next-to-last term, then substituting $[H_3O^+]$ for K_{n-1} to obtain the next term, etc, until the first term is reached. The following examples show the denominator, D, to be used for various types of acids

$$H_3A\text{:}\quad D = [H_3O^+]^3 + K_1[H_3O^+]^2 + K_1K_2[H_3O^+] + K_1K_2K_3 \tag{49}$$

$$H_2A^-\text{:}\quad D = [H_3O^+]^2 + K_1[H_3O^+] + K_1K_2 \tag{50}$$

$$HA^{2-}\text{:}\quad D = [H_3O^+] + K_a \tag{51}$$

The numerator, in all instances, is C_a multiplied by the term from the denominator that has $[H_3O^+]$ raised to the $n - j$ power. Thus, for diprotic acids such as carbonic, succinic, tartaric, etc

$$[H_2A] = \frac{[H_3O^+]^2 C_a}{[H_3O^+]^2 + K_1[H_3O^+] + K_1K_2} \tag{52}$$

$$[HA^-] = \frac{K_1[H_3O^+] C_a}{[H_3O^+]^2 + K_1[H_3O^+] + K_1K_2} \tag{53}$$

$$[A^{2-}] = \frac{K_1K_2 C_a}{[H_3O^+]^2 + K_1[H_3O^+] + K_1K_2} \tag{54}$$

Example—Calculate the concentrations of all succinic acid species in a 1.0×10^{-3} *M* solution of succinic acid at pH 6.0. Assume that $K_1 = 6.4 \times 10^{-5}$ and $K_2 = 2.3 \times 10^{-6}$.

Eqs 52–54 have the same denominator, D, which can be calculated as

$$D = [H_3O^+]^2 + K_1[H_3O^+] + K_1K_2$$

$$= 1.0 \times 10^{-12} + 6.4 \times 10^{-5} \times 1.0 \times 10^{-6} + 6.4 \times 10^{-5} \times 2.3 \times 10^{-6}$$

$$= 1.0 \times 10^{-12} + 6.4 \times 10^{-11} + 14.7 \times 10^{-11}$$

$$= 21.2 \times 10^{-11}$$

Therefore

$$[H_2A] = \frac{[H_3O^+]^2 C_a}{D} = \frac{1.0 \times 10^{-12} \times 1.0 \times 10^{-3}}{21.2 \times 10^{-11}} = 4.7 \times 10^{-6}\ M$$

* The pH of the purest water obtainable, so-called "conductivity" water, is 7.0 when the measurement is carefully made under conditions to exclude carbon dioxide and prevent errors inherent in the measuring technique (such as acidity or alkalinity of the indicator). Upon agitating this water in the presence of carbon dioxide in the atmosphere (equilibrium water) the value drops rapidly to 5.7, which is the pH of nearly all distilled waters that have been exposed to the atmosphere for even a short time.

$$[HA^-] = \frac{K_1[H_3O^+]C_a}{D} =$$

$$\frac{6.4 \times 10^{-11} \times 1.0 \times 10^{-3}}{21.2 \times 10^{-11}} = 3.0 \times 10^{-4}\ M$$

$$[A^{2-}] = \frac{K_1K_2C_a}{D} =$$

$$\frac{14.7 \times 10^{-11} \times 1.0 \times 10^{-3}}{21.2 \times 10^{-11}} = 6.9 \times 10^{-4}\ M$$

Proton Balance Equation

In the Brønsted–Lowry system the total number of protons released by acidic species must equal the total number of protons consumed by basic species. This results in a very useful relationship known as the proton balance equation (PBE), in which the sum of the concentration terms for species that form by proton consumption is equated to the sum of the concentration terms for species that are formed by the release of protons. The PBE forms the basis of a unified approach to pH calculations, since it is an exact accounting of all proton transfers occurring in solution.

When HCl is added to water, for example, it dissociates yielding one Cl^- for each proton released. Thus, Cl^- is a species formed by the release of a proton. In the same solution, and actually in all aqueous solutions

$$2H_2O \rightleftharpoons H_3O^+ + OH^-$$

where H_3O^+ is formed by proton consumption and OH^- is formed by proton release. Thus, the PBE is

$$[H_3O^+] = [OH^-] + [Cl^-] \tag{55}$$

In general, the PBE can be formed in the following manner

1. Start with the species added to water.
2. Place all species that can form when protons are released on the right side of the equation.
3. Place all species that can form when protons are consumed on the left side of the equation.
4. Add $[H_3O^+]$ to the left side of the equation and $[OH^-]$ to the right side of the equation. These result from the interaction of two molecules of water as shown above.

Example—When H_3PO_4 is added to water, the species $H_2PO_4^-$ forms with the release of one proton, HPO_4^{2-} forms with the release of two protons and PO_4^{3-} forms with the release of three protons to give the following PBE

$$[H_3O^+] = [OH^-] + [H_2PO_4^-] + 2[HPO_4^{2-}] + 3[PO_4^{3-}] \tag{56}$$

Example—When Na_2HPO_4 is added to water, it dissociates into two Na^+ and one HPO_4^{2-}. The sodium ion is neglected in the PBE since it is not formed from the release or consumption of protons. The species HPO_4^{2-}, however, may react with water to give $H_2PO_4^-$ with the consumption of one proton, H_3PO_4 with the consumption of two protons, and PO_4^{3-} with the release of one proton to give the following PBE

$$[H_3O^+] + [H_2PO_4^-] + 2[H_3PO_4] = [OH^-] + [PO_4^{3-}] \tag{57}$$

Calculations

The pH of solutions of acids, bases and salts may be calculated using the concepts presented in the preceding sections.

Strong Acids or Bases

When a strong acid such as HCl is added to water, the following reactions occur

$$HCl + H_2O \rightarrow H_3O^+ + Cl^-$$

$$2H_2O \rightleftharpoons H_3O^+ + OH^-$$

The proton balance equation for this system would be

$$[H_3O^+] = [OH^-] + [Cl^-] \tag{58}$$

In most instances ($C_a > 4.5 \times 10^{-7}\ M$) the $[OH^-]$ would be negligible compared to the $[Cl^-]$ and the equation simplifies to

$$[H_3O^+] = [Cl^-] = C_a \tag{59}$$

Thus, the hydronium-ion concentration of a solution of a strong acid would be equal to the stoichiometric concentration of the acid. This would be anticipated, since strong acids generally are assumed to be 100% ionized.

The pH of a 0.005 M solution of HCl therefore is calculated as

$$pH = -\log 0.005 = 2.30$$

In a similar manner the hydroxyl-ion concentration for a solution of a strong base such as NaOH would be

$$[OH^-] = [Na^+] = C_b \tag{60}$$

and the pH of a 0.005 M solution of NaOH would be

$$pOH = -\log 0.005 = 2.30$$

$$pH = pK_w - pOH = 14.00 - 2.30 = 11.70$$

Weak Acids or Bases

If a weak acid, HA, is added to water, it will equilibrate with its conjugate base, A^-, as

$$HA + H_2O \rightleftharpoons H_3O^+ + A^-$$

Accounting for the ionization of water gives the following proton balance equation for this system

$$[H_3O^+] = [OH^-] + [A^-] \tag{61}$$

The concentration of A^- as a function of hydronium-ion concentration can be obtained as shown previously to give

$$[H_3O^+] = [OH^-] + \frac{K_aC_a}{[H_3O^+] + K_a} \tag{62}$$

Algebraic simplification yields

$$[H_3O^+] = K_a \frac{(C_a - [H_3O^+] + [OH^-])}{([H_3O^+] - [OH^-])} \tag{63}$$

In most instances for solutions of weak acids, $[H_3O^+] \gg [OH^-]$ and the equation simplifies to give

$$[H_3O^+]^2 + K_a[H_3O^+] - K_aC_a = 0 \tag{64}$$

This is a quadratic equation* which yields

$$[H_3O^+] = \frac{-K_a + \sqrt{K_a^2 + 4K_aC_a}}{2} \tag{65}$$

since $[H_3O^+]$ can never be negative. Furthermore, if $[H_3O^+]$ is less than 5% of C_a, Eq 64 is simplified further to give

* The general solution to a quadratic equation of the form

$$aX^2 + bX + c = 0$$

is

$$X = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$[H_3O^+] = \sqrt{K_a C_a} \tag{66}$$

It generally is preferable to use the simplest equation to calculate $[H_3O^+]$. However, when $[H_3O^+]$ is calculated, it must be compared to C_a in order to determine whether the assumption $C_a \gg [H_3O^+]$ is valid. If the assumption is not valid, the quadratic equation should be used.

Example—Calculate the pH of a 5.00×10^{-5} *M* solution of a weak acid having a $K_a = 1.90 \times 10^{-5}$.

$$[H_3O^+] = \sqrt{K_a C_a}$$
$$= 1.90 \times 10^{-5} \times 5.00 \times 10^{-5}$$
$$= 3.08 \times 10^{-5}\ M$$

Since C_a [$(5.00 \times 10^{-5}\ M)$] is not much greater than $[H_3O^+]$, the quadratic equation (Eq 65) should be used

$$[H_3O^+] = \frac{-1.90 \times 10^{-5} + \sqrt{(1.90 \times 10^{-5})^2 + 4(1.90 \times 10^{-5} \times 5.00 \times 10^{-5})}}{2}$$
$$= 2.26 \times 10^{-5}\ M$$
$$\text{pH} = -\log (2.26 \times 10^{-5}) = 4.65$$

Note that the assumption $[H_3O^+] \gg [OH^-]$ is valid. The hydronium-ion concentration calculated from Eq 66 has a relative error of 36% when compared to the correct value obtained from Eq 65.

When a salt obtained from a strong acid and a weak base—e.g., ammonium chloride, morphine sulfate, pilocarpine hydrochloride, etc—is dissolved in water, it dissociates as

$$BH^+X^- \xrightarrow{H_2O} BH^+ + X^-$$

in which BH^+ is the protonated form of the base B, and X^- is the anion of a strong acid. Since X^- is the anion of a strong acid, it is too weak a base to undergo any further reaction with water. The protonated base, however, can act as a weak acid to give

$$BH^+ + H_2O \rightleftharpoons B + H_3O^+$$

Thus, Eqs 65 and 66 are valid, with C_a being equal to the concentration of the salt in solution. If K_a for the protonated base is not available, it can be obtained by dividing K_b for the base B, into K_w.

Example—Calculate the pH of a 0.026 *M* solution of ammonium chloride. Assume that K_b for ammonia is 1.74×10^{-5} and K_w is 1.00×10^{-14}.

$$K_a = \frac{K_w}{K_b} = \frac{1.00 \times 10^{-14}}{1.74 \times 10^{-5}} = 5.75 \times 10^{-10}$$
$$[H_3O^+] = \sqrt{K_a C_a}$$
$$= \sqrt{5.75 \times 10^{-10} \times 2.6 \times 10^{-2}}$$
$$= 3.87 \times 10^{-6}\ M$$
$$\text{pH} = -\log (3.87 \times 10^{-6}) = 5.41$$

Since C_a is much greater than $[H_3O^+]$ and $[H_3O^+]$ is much greater than $[OH^-]$, the assumptions are valid and the value calculated for pH is sufficiently accurate.

Weak Bases

When a weak base, B, is dissolved in water it ionizes to give the conjugate acid as

$$B + H_2O \rightleftharpoons BH^+ + OH^-$$

The proton balance equation for this system is

$$[BH^+] + [H_3O^+] = [OH^-] \tag{67}$$

Substituting $[BH^+]$ as a function of hydronium-ion concentration and simplifying, in the same manner as shown for a weak acid, gives

$$[OH^-] = K_b \frac{(C_b - [OH^-] + [H_3O^+])}{([OH^-] - [H_3O^+])} \tag{68}$$

If $[OH^-] \gg [H_3O^+]$, as is true generally

$$[OH^-]^2 + K_b[OH^-] - K_b C_b = 0 \tag{69}$$

which is a quadratic with the following solution

$$[OH^-] = \frac{-K_b + \sqrt{K_b^2 + 4K_b C_b}}{2} \tag{70}$$

If $C_b \gg [OH^-]$, the quadratic equation simplifies to

$$[OH^-] = \sqrt{K_b C_b} \tag{71}$$

Once $[OH^-]$ is calculated, it can he converted to pOH, which can be subtracted from pK_w to give pH.

Example—Calculate the pH of a 4.50×10^{-2} *M* solution of a weak base having $K_b = 2.00 \times 10^{-4}$. Assume that $K_w = 1.00 \times 10^{-14}$.

$$[OH^-] = \sqrt{K_b C_b}$$
$$= \sqrt{2.00 \times 10^{-4} \times 4.50 \times 10^{-2}}$$
$$= \sqrt{9.00 \times 10^{-6}} = 3.00 \times 10^{-3}\ M$$

Both assumptions are valid.

$$\text{pOH} = -\log 3.00 \times 10^{-3} = 2.52$$
$$\text{pH} = 14.00 - 2.52 = 11.48$$

When salts obtained from strong bases and weak acids (eg, sodium acetate, sodium sulfathiazole, sodium benzoate, etc) are dissolved in water, they dissociate as

$$Na^+A^- \xrightarrow{H_2O} Na^+ + A^-$$

in which A^- is the conjugate base of the weak acid, HA. The Na^+ undergoes no further reaction with water. The A^-, however, acts as a weak base to give

$$A^- + H_2O \rightleftharpoons HA + OH^-$$

Thus, Eqs 70 and 71 are valid, with C_b being equal to the concentration of the salt in solution. The value for K_b can be obtained by dividing K_a for the conjugate acid, HA, into K_w.

Example—Calculate the pH of a 0.05 *M* solution of sodium acetate. Assume that K_a for acetic acid = 1.75×10^{-5} and $K_w = 1.00 \times 10^{-14}$.

$$K_b = \frac{K_w}{K_a} = \frac{1.00 \times 10^{-14}}{1.75 \times 10^{-5}}$$
$$= 5.71 \times 10^{-10}$$
$$OH^- = \sqrt{K_b C_b} = \sqrt{5.71 \times 10^{-10} \times 5.0 \times 10^{-2}}$$
$$= 5.34 \times 10^{-6}\ M$$

Both assumptions are valid

$$\text{pOH} = -\log (5.34 \times 10^{-6}) = 5.27$$
$$\text{pH} = 14.00 - 5.27 = 8.73$$

Ampholytes

Substances such as $NaHCO_3$ and NaH_2PO_4 are termed *ampholytes*, and are capable of functioning both as acids and bases. When an ampholyte of the type NaHA is dissolved in water, the following series of reactions can occur

$$Na^+HA^- \xrightarrow{H_2O} Na^+ + HA^-$$

$$HA^- + H_2O \rightleftharpoons A^{2-} + H_3O^+$$

$$HA^- + H_2O \rightleftharpoons H_2A + OH^-$$

$$2H_2O \rightleftharpoons H_3O^+ + OH^-$$

The total proton balance equation (PBE) for the system is

$$[H_3O^+] + [H_2A] = [OH^-] + [A^{2-}] \tag{72}$$

Substituting both $[H_2A]$ and $[A^{2-}]$ as a function of $[H_3O^+]$ (see Eqs 52 and 54), yields

$$[H_3O^+] + \frac{[H_3O^+]^2 C_s}{[H_3O^+]^2 + K_1[H_3O^+] + K_1K_2} = \frac{K_w}{[H_3O^+]} + \frac{K_1K_2C_s}{[H_3O^+]^2 + K_1[H_3O^+] + K_1K_2}$$

This gives a fourth-order equation in $[H_3O^+]$, which can be simplified using certain judicious assumptions to

$$[H_3O^+] = \sqrt{\frac{K_1K_2C_s}{K_1 + C_s}} \tag{73}$$

In most instances, $C_s \gg K_1$ and the equation further simplifies to

$$[H_3O^+] = \sqrt{K_1K_2} \tag{74}$$

and $[H_3O^+]$ becomes independent of the concentration of the salt. A special property of ampholytes is that the concentration of the species HA^- is maximum at the pH corresponding to Eq 74.

When the simplest amino acid salt, glycine hydrochloride, is dissolved in water, it acts as a diprotic acid and ionizes as

$$^+NH_3CH_2COOH + H_2O \rightleftharpoons {}^+NH_3CH_2COO^- + H_3O^+$$

$$^+NH_3CH_2COO^- + H_2O \rightleftharpoons NH_2CH_2COO^- + H_3O^+$$

The form, $^+NH_3CH_2COO^-$, is an ampholyte since it also can act as a weak base

$$^+NH_3CH_2COO^- + H_2O \rightleftharpoons {}^+NH_3CH_2COOH + OH^-$$

This type of substance, which carries both a charged acidic and a charged basic moiety on the same molecule is termed a *zwitterion* and, since the two charges balance each other, the molecule acts essentially as a neutral molecule. The pH at which the *zwitterion* concentration is maximum is known as the isoelectric point, which can be calculated from Eq 74.

On the acid side of the isoelectric point, amino acids and proteins are cationic and incompatible with anionic materials such as the naturally occurring gums used as suspending and/or emulsifying agents. On the alkaline side of the isoelectric point, amino acids and proteins are anionic and incompatible with cationic materials such as benzalkonium chloride.

Salts of Weak Acids and Weak Bases

When a salt such as ammonium acetate (which is derived from a weak acid and a weak base) is dissolved in water, it undergoes the following reactions

$$BH^+A^- \xrightarrow{H_2O} BH^+ + A^-$$

$$BH^+ + H_2O \rightleftharpoons B + H_3O^+$$

$$A^- + H_2O \rightleftharpoons HA + OH^-$$

The total PBE for this system is

$$[H_3O^+] + [HA] = [OH^-] + [B] \tag{75}$$

Replacing [HA] and [B] as a function of $[H_3O^+]$, gives

$$[H_3O^+] + \frac{[H_3O^+]C_s}{[H_3O^+] + K_a} = [OH^-] + \frac{K_a'C_s}{[H_3O^+] + K_a'} \tag{76}$$

in which C_s is the concentration of salt, K_a is the ionization constant of the conjugate acid formed from the reaction between A^- and water and K_a' is the ionization constant for the protonated base, BH^+. In general, $[H_3O^+]$, $[OH^-]$, K_a and K_a' usually are smaller than C_s and the equation simplifies to

$$[H_3O^+] = \sqrt{K_aK_a'} \tag{77}$$

Example—Calculate the pH of a 0.01 *M* solution of ammonium acetate. The ammonium ion has a K_a equal to 5.75×10^{-10}, which represents K_a' in Eq 77. Acetic acid has a K_a of 1.75×10^{-5}, which represents K_a in Eq 77

$$[H_3O^+] = \sqrt{1.75 \times 10^{-5} \times 5.75 \times 10^{-10}}$$

$$= 1.05 \times 10^{-7}$$

$$pH = -\log (1.05 \times 10^{-7}) = 6.98$$

All of the assumptions are valid.

Buffers

The terms *buffer*, *buffer solution* and *buffered solution*, when used with reference to hydrogen-ion concentration or pH, refer to the ability of a system, particularly an aqueous solution, to resist a change of pH on adding acid or alkali, or on dilution with a solvent.

If an acid or base is added to water, the pH of the latter is changed markedly, for water has no ability to resist change of pH; it is completely devoid of buffer action. Even a very weak acid such as carbon dioxide changes the pH of water, decreasing it from 7 to 5.7 when the small concentration of carbon dioxide present in air is equilibrated with pure water. This extreme susceptibility of distilled water to a change of pH on adding very small amounts of acid or base is often of great concern in pharmaceutical operations. Solutions of neutral salts, such as sodium chloride, similarly lack ability to resist change of pH on adding acid or base; such solutions are called unbuffered.

Characteristic of buffered solutions, which undergo small changes of pH on addition of acid or base, is the presence either of a weak acid and a salt of the weak acid, or a weak base and a salt of the weak base. An example of the former system is acetic acid and sodium acetate; of the latter, ammonium hydroxide and ammonium chloride. From the proton concept of acids and bases discussed earlier, it is apparent that such buffer action involves a conjugate acid–base pair in the solution. It will be recalled that acetate ion is the conjugate base of acetic acid, and that ammonium ion is the conjugate acid of ammonia (the principal constituent of what commonly is called ammonium hydroxide).

The mechanism of action of the acetic acid–sodium acetate buffer pair is that the acid, which exists largely in molecular (nonionized) form, combines with hydroxyl ion that may be added to form acetate ion and water, thus

$$CH_3COOH + OH^- \rightarrow CH_3COO^- + H_2O$$

while the acetate ion, which is a base, combines with hydrogen (more exactly hydronium) ion that may be added to

form essentially nonionized acetic acid and water, represented as

$$CH_3COO^- + H_3O^+ \rightarrow CH_3COOH + H_2O$$

As will be illustrated later by an example, the change of pH is slight as long as the amount of hydronium or hydroxyl ion added does not exceed the capacity of the buffer system to neutralize it.

The ammonia–ammonium chloride pair functions as a buffer because the ammonia combines with hydronium ion that may be added to form ammonium ion and water, thus

$$NH_3 + H_3O^+ \rightarrow NH_4^+ + H_2O$$

Ammonium ion, which is an acid, combines with added hydroxyl ion to form ammonia and water, as

$$NH_4^+ + OH^- \rightarrow NH_3 + H_2O$$

Again, the change of pH is slight if the amount of added hydronium or hydroxyl ion is not in excess of the capacity of the system to neutralize it.

Besides these two general types of buffers, a third appears to exist. This is the buffer system composed of two salts, as monobasic potassium phosphate, KH_2PO_4, and dibasic potassium phosphate, K_2HPO_4. This is not, however, a new type of buffer; it is actually a weak-acid–conjugate-base buffer in which an ion, $H_2PO_4^-$, serves as the weak acid, and HPO_4^{2-} is its conjugate base. When hydroxyl ion is added to this buffer the following reaction takes place

$$H_2PO_4^- + OH^- \rightarrow HPO_4^{2-} + H_2O$$

and when hydronium ion is added

$$HPO_4^{2-} + H_3O^+ \rightarrow H_2PO_4^- + H_2O$$

It is apparent that the mechanism of action of this type of buffer is essentially the same as that of the weak-acid–conjugate-base buffer composed of acetic acid and sodium acetate.

Calculations—A buffer system composed of a conjugate acid–base pair, NaA – HA (such as sodium acetate and acetic acid), would have a PBE of

$$[H_3O^+] + [HA] = [OH^-] + [A^-] \tag{78}$$

Replacing [HA] and $[A^-]$ as a function of hydronium-ion concentration gives

$$[H_3O^+] + \frac{[H_3O^+]C_b}{[H_3O^+] + K_a} = [OH^-] + \frac{K_aC_a}{[H_3O^+] + K_a} \tag{79}$$

where C_b is the concentration of the salt, NaA, and C_a is the concentration of the weak acid, HA. This equation can be rearranged to give

$$[H_3O^+] = K_a \frac{(C_a - [H_3O^+] + [OH^-])}{(C_b + [H_3O^+] - [OH^-])} \tag{80}$$

In general, both C_a and C_b are much greater than $[H_3O^+]$, which is in turn much greater than $[OH^-]$, and the equation simplifies to

$$[H_3O^+] = \frac{K_aC_a}{C_b} \tag{81}$$

or, expressed in terms of pH, as

$$pH = pK_a + \log \frac{C_b}{C_a} \tag{82}$$

This equation is generally called the Henderson–Hasselbalch equation. It applies to all buffer systems formed from a single conjugate acid–base pair, regardless of the nature of the salts. For example, it applies equally well to the following buffer systems: ammonia–ammonium chloride, monosodium phosphate–disodium phosphate, phenobarbital–sodium phenobarbital, etc. In the ammonia–ammonium chloride system, ammonia is obviously the base and the ammonium ion is the acid (C_a equal to the concentration of the salt). In the phosphate system, monosodium phosphate is the acid and disodium phosphate is the base. For the phenobarbital buffer system, phenobarbital is the acid and the phenobarbital anion is the base (C_b equal to the concentration of sodium phenobarbital).

As an example of the application of this equation, the pH of a buffer solution containing acetic acid and sodium acetate, each in 0.1 *M* concentration, may be calculated. The K_a of acetic acid, as defined above, is 1.8×10^{-5}, at 25°.

Solution

First, the pK_a of acetic acid is calculated

$$pK_a = -\log K_a = -\log 1.8 \times 10^{-5}$$
$$= -\log 1.8 - \log 10^{-5}$$
$$= -0.26 - (-5) = +4.74$$

Substituting this value into Eq 82

$$pH = \log \frac{0.1}{0.1} + 4.74 = +4.74$$

The Henderson–Hasselbalch equation predicts that any solutions containing the same molar concentration of acetic acid as of sodium acetate will have the same pH. Thus, a solution of 0.01 *M* concentration of each will have the same pH, 4.74, as one of 0.1 *M* concentration of each component. Actually, there will be some difference in the pH of the solutions, for the *activity coefficient* of the components varies with concentration. For most practical purposes, however, the approximate values of pH calculated by the equation are satisfactory. It should be pointed out, however, that the buffer of higher concentration of each component will have a much greater capacity for neutralizing added acid or base and this point will be discussed further under *Buffer Capacity*.

The Henderson–Hasselbalch equation is useful also for calculating the ratio of molar concentrations of a buffer system required to produce a solution of specific pH. As an example, suppose that an acetic acid–sodium acetate buffer of pH 4.5 is to be prepared. What ratio of the buffer components should be used?

Solution

Rearranging Eq 82, which is used to calculate the pH of weak acid-salt type buffers, gives

$$\log \frac{[\text{base}]}{[\text{acid}]} = pH - pK_a$$
$$= 4.5 - 4.76 = -0.24 = (9.76 - 10)$$
$$\frac{[\text{base}]}{[\text{acid}]} = \text{antilog of } (9.76 - 10) = 0.575$$

The interpretation of this result is that the *proportion* of sodium acetate to acetic acid should be 0.575 mole of the former to 1 mole of the latter to produce a pH of 4.5. A solution containing 0.0575 mole of sodium acetate and 0.1 mole of acetic acid per liter would meet this requirement, as would also one containing 0.00575 mole of sodium acetate and 0.01 mole of acetic acid per liter. The actual concentration selected would depend chiefly on the desired buffer capacity.

Buffer Capacity—The ability of a buffer solution to resist changes in pH upon addition of acid or alkali may be measured in terms of *buffer capacity*. In the preceding discussion of buffers, it has been seen that, in a general way, the concentration of acid in a weak-acid–conjugate-base buffer determines the capacity to "neutralize" added base,

while the concentration of salt of the weak acid determines the capacity to neutralize added acid. Similarly, in a weak-base–conjugate-acid buffer the concentration of the weak base establishes the buffer capacity toward added acid, while the concentration of the conjugate acid of the weak base determines the capacity toward added base. When the buffer is equimolar in the concentrations of weak acid and conjugate base, or of weak base and conjugate acid, it has equal buffer capacity toward added strong acid or strong base.

Van Slyke, the biochemist, introduced a quantitative expression for evaluating buffer capacity. This may be defined as the amount, in gram-equivalents (g-Eq) per liter, of strong acid or strong base, required to be added to a solution to change its pH by 1 unit; a solution has a buffer capacity of 1 when 1 L requires 1 g-Eq of strong base or acid to change the pH 1 unit (in practice, considerably smaller increments are measured, expressed as the ratio of acid or base added to the change of pH produced). From this definition it is apparent that the smaller the pH change in a solution caused by the addition of a specified quantity of acid or alkali, the greater the buffer capacity of the solution.

The following numerical examples illustrate certain basic principles and calculations concerning buffer action and buffer capacity.

Example 1—What is the change of pH on adding 0.01 mole of NaOH to 1 L of 0.10 *M* acetic acid?

(a) Calculate the pH of a 0.10 molar solution of acetic acid

$$[H_3O^+] = \sqrt{K_a C_a} = \sqrt{1.75 \times 10^{-4} \times 1.0 \times 10^{-1}} = 1.33 \times 10^{-3}$$

$$pH = -\log 1.33 \times 10^{-3} = 2.88$$

(b) On adding 0.01 mole of NaOH to a liter of this solution, 0.01 mole of acetic acid is converted to 0.01 mole of sodium acetate, thereby decreasing C_a to 0.09 *M*, and $C_b = 1.0 \times 10^{-2}$ *M*. Using the Henderson–Hasselbach equation gives

$$pH = 4.76 + \log \frac{0.01}{0.09} = 4.76 - 0.95 = 3.81$$

The pH change is, therefore, 0.93 unit. The buffer capacity as defined above is calculated to be

$$\frac{\text{moles of NaOH added}}{\text{change in pH}} = 0.011$$

Example 2—What is the change of pH on adding 0.1 mole of NaOH to 1 L of buffer solution 0.1 *M* in acetic acid and 0.1 *M* in sodium acetate?

(a) The pH of the buffer solution before adding NaOH is

$$pH = \log \frac{[\text{base}]}{[\text{acid}]} + pK_a$$

$$= \log \frac{0.1}{0.1} + 4.76 = 4.76$$

(b) On adding 0.01 mole of NaOH per liter to this buffer solution, 0.01 mole of acetic acid is converted to 0.01 mole of sodium acetate, thereby decreasing the concentration of acid to 0.09 *M* and increasing the concentration of base to 0.11 *M*. The pH is calculated as

$$pH = \log \frac{0.11}{0.09} + 4.76$$

$$= 0.086 + 4.76 = 4.85$$

The change of pH in this case is only 0.09 unit, about 1/10 the change in the preceding example. The buffer capacity is calculated as

$$\frac{\text{moles of NaOH added}}{\text{change of pH}} = \frac{0.01}{0.09} = 0.11$$

Thus, the buffer capacity of the acetic acid–sodium acetate buffer solution is approximately 10 times that of the acetic acid solution.

As is in part evident from these examples, and may be further evidenced by calculations of pH changes in other systems, the degree of buffer action and, therefore, the buffer capacity, depend on the kind and concentration of the buffer components, the pH region involved and the kind of acid or alkali added.

Strong Acids and Bases as "Buffers"—In the foregoing discussion, buffer action was attributed to systems of (1) weak acids and their conjugate bases, (2) weak bases and their conjugate acids and (3) certain acid-base pairs which can function in the manner either of System 1 or 2.

The ability to resist change in pH on adding acid or alkali is possessed also by relatively concentrated solutions of strong acids and strong bases. If to 1 L of pure water having a pH of 7.0 is added 1 mL of 0.01 *M* hydrochloric acid, the pH is reduced to about 5.0. If the same volume of the acid is added to 1 L of 0.001 *M* hydrochloric acid, which has a pH of about 3, the hydronium-ion concentration is increased only about 1% and the pH is reduced hardly at all. The nature of this buffer action is quite different from that of the true buffer solutions. The very simple explanation is that when 1 mL of 0.01 *M* HCl, which represents 0.00001 g-Eq of hydronium ions, is added to the 0.0000001 g-Eq of hydronium ions in 1 L of pure water, the hydronium-ion concentration is increased 100-fold (equivalent to 2 pH units), but when the same amount is added to the 0.001 g-Eq of hydronium ions in 1 L of 0.001 *M* HCl, the increase is only 1/100 the concentration already present. Similarly, if 1 mL of 0.01 *M* NaOH is added to 1 L of pure water, the pH is increased to 9, while if the same volume is added to 1 L of 0.001 molar NaOH, the pH is increased almost immeasurably.

In general, solutions of strong acids of pH 3 or less, and solutions of strong bases of pH 11 or more, exhibit this kind of buffer action by virtue of the relatively high concentration of hydronium or hydroxyl ions present. The USP includes among its *Standard Buffer Solutions* a series of hydrochloric acid buffers, covering the pH range 1.2 to 2.2, which also contain potassium chloride. The salt does not participate in the buffering mechanism, as is the case with salts of weak acids; instead, it serves as a nonreactive constituent required to maintain the proper electrolyte environment of the solutions.

Determination of pH

Colorimetry

A relatively simple and inexpensive method for determining the approximate pH of a solution depends on the fact that some conjugate acid base pairs (indicators) possess one color in the acid form and another color in the base form. Assume that the acid form of a particular indicator is red, while the base form is yellow. The color of a solution of this indicator will range from red, when it is sufficiently acid, to yellow, when it is sufficiently alkaline. In the intermediate pH range (the transition interval) the color will be a blend of red and yellow depending upon the ratio of the base to the acid form. In general, although there are slight differences between indicators, color changes apparent to the eye cannot be discerned when the ratio of base to acid form, or acid to base form exceeds 10:1. The use of Eq 82 indicates that the transition range of most indicators is equal to the pK_a of the indicator ±1 pH unit, or a useful range of approximately 2 pH units. Standard indicator solutions can be made at known pH values within the transition range of the indicator, and the pH of an unknown solution determined by adding the indicator to it and comparing the resulting color with the standard solutions. Details of this procedure can be found in RPS-14. Another method for using these indicators is to apply them to thin strips of filter paper. A drop of the unknown solution is placed on a piece of the indicator paper and the resulting color compared to a color chart supplied with the indicator paper. These papers are available in a wide variety of pH ranges.

Potentiometry

Electrometric methods for the determination of pH are based on the fact that the difference of electrical potential between two suitable electrodes dipping into a solution containing hydronium ions depends on the concentration (or activity) of the latter. The development of a potential difference is not a specific property of hydronium ions. A solution of any ion will develop a potential proportional to the concentration of that ion if a suitable pair of electrodes is placed in the solution.

The relationship between the potential difference and concentration of an ion in equilibrium with the electrodes may be derived as follows. When a metal is immersed into a solution of one of its salts, there is a tendency for the metal to go into solution in the form of ions. This tendency is known as the *solution pressure* of the metal and is comparable to the tendency of sugar molecules, for example, to dissolve in water. The metallic ions in solution tend, on the other hand, to become discharged by forming atoms, this effect being proportional to the *osmotic pressure* of the ions. In order for an atom of a metal to go into solution as a positive ion, electrons, equal in number to the charge on the ion, must be left behind on the metal electrode with the result that the latter becomes negatively charged. The positively charged ions in solution, however, may become discharged as atoms by taking up electrons from the metal electrode. Depending on which effect predominates, the electrical charge on the electrode will be either positive or negative and may be quantitatively expressed by the following equation proposed by Nernst in 1889

$$E = \frac{RT}{nF} \ln \frac{p}{P} \tag{83}$$

where E is the potential difference or electromotive force, R is the gas constant (8.316 joules), T is the absolute temperature, n is the valence of the ion, F is the Faraday of electricity (96,500 coulombs), p is the osmotic pressure of the ions and P is the solution pressure of the metal.

Inasmuch as it is impossible to measure the potential difference between one electrode and a solution with any degree of certainty, it is customary to use two electrodes and to measure the potential difference between them. If two electrodes, both of the same metal, are immersed separately in solutions containing ions of that metal, at osmotic pressure p_1 and p_2, respectively, and connected by means of a tube containing a nonreacting salt solution (a so-called "salt-bridge"), the potential developed across the two electrodes will be equal to the difference between the potential differences of the individual electrodes; thus

$$E = E_1 - E_2 = \frac{RT}{nF} \ln \frac{p_1}{P_1} - \frac{RT}{nF} \ln \frac{p_2}{P_2} \tag{84}$$

Since both electrodes are of the same metal, $P_1 = P_2$ and the equation may be simplified to

$$E = \frac{RT}{nF} \ln p_1 - \frac{RT}{nF} \ln p_2 = \frac{RT}{nF} \ln \frac{p_1}{p_2} \tag{85}$$

In place of osmotic pressures it is permissible, for dilute solutions, to substitute the concentrations c_1 and c_2 which were found (see Chapter 16, page 222) to be proportional to p_1 and p_2. The equation then becomes

$$E = \frac{RT}{nF} \ln \frac{c_1}{c_2} \tag{86}$$

If either c_1 or c_2 is known, it is obvious that the value of the other may be found if the potential difference, E, of this cell can be measured.

For the determination of hydronium-ion concentration or pH, an electrode at which an equilibrium between hydrogen gas and hydronium ion can be established must be used in place of metallic electrodes. Such an electrode may be made by electrolytically coating a strip of platinum, or other noble metal, with platinum black and saturating the latter with pure hydrogen gas. This device functions as a *hydrogen electrode*. Two such electrodes may be assembled as shown in Fig 17-3.

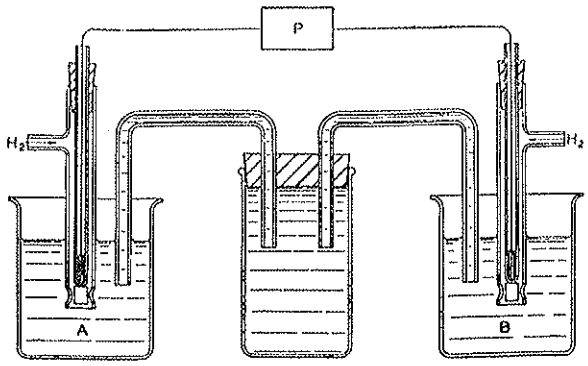


Fig 17-3. Hydrogen-ion concentration chain.

In this diagram one electrode dips into Solution A, containing a known hydronium-ion concentration, and the other electrode dips into Solution B, containing an unknown hydronium-ion concentration. The two electrodes and solutions, sometimes called half-cells, then are connected by a bridge of neutral salt solution, which has no significant effect on the solutions it connects. The potential difference across the two electrodes is measured by means of a potentiometer, P. If the concentration, c_1, of hydronium ion in Solution A is 1 N, Eq 86 simplifies to

$$E = \frac{RT}{nF} \ln \frac{1}{c_2} \tag{87}$$

or in terms of Briggsian logarithms

$$E = 2.303 \frac{RT}{nF} \log_{10} \frac{1}{c_2} \tag{88}$$

If for $\log_{10} 1/c_2$ there is substituted its equivalent pH, the equation becomes

$$E = 2.303 \frac{RT}{nF} \text{ pH} \tag{89}$$

and finally by substituting numerical values for R, n, T and F, and assuming the temperature to be 20°, the following simple relationship is derived

$$E = 0.0581 \text{ pH or pH} = \frac{E}{0.0581} \tag{90}$$

The hydrogen electrode dipping into a solution of known hydronium-ion concentration, called the *reference electrode*, may be replaced by a calomel electrode, one type of which is shown in Fig 17-4. The elements of a calomel electrode are mercury and calomel in an aqueous solution of potassium chloride. The potential of this electrode is constant, regardless of the hydronium-ion concentration of the solution into which it dips. The potential depends on the equilibrium which is set up between mercury and mercurous ions from the calomel, but the concentration of the latter is governed, according to the solubility-product principle, by the concentration of chloride ions, which are derived mainly from the potassium chloride in the solution. Therefore, the potential of this electrode varies with the concentration of potassium chloride in the electrolyte.

Because the calomel electrode always indicates voltages which are higher, by a constant value, than those obtained

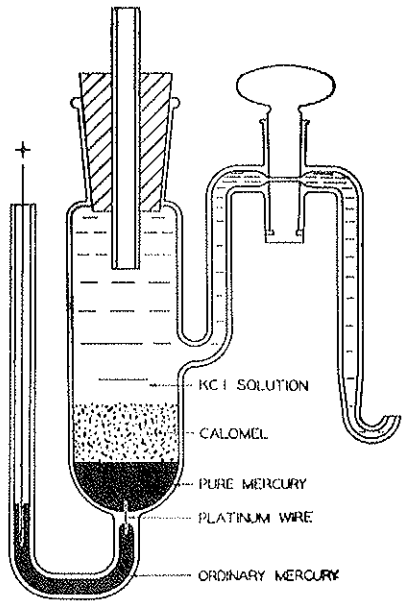


Fig 17-4. Calomel electrode.

when the normal hydrogen electrode chain shown in Fig 17-3 is used, it is necessary to subtract the potential due to the calomel electrode itself from the observed voltage. As the magnitude of this voltage depends on the concentration of potassium chloride in the calomel-electrode electrolyte, it is necessary to know the concentration of the former. For most purposes a saturated potassium chloride solution is used which produces potential difference of 0.2488 v. Accordingly, before using Eq 85 for the calculation of pH from the voltage of a cell made up of a calomel and a hydrogen electrode dipping into the solution to be tested, 0.2488 v must be subtracted from the observed potential difference. Expressed mathematically, Eq 91 is used for calculating pH from the potential difference of such a cell.

$$\text{pH} = \frac{E - 0.2488}{0.0581} \tag{91}$$

In measuring the potential difference between the electrodes, it is imperative that very little current be drawn from the cell, for with current flowing the voltage changes, owing to polarization effects at the electrode. Because of this it is not possible to make accurate measurements with a voltmeter which requires appreciable current to operate it. In its place a potentiometer is used which does not draw a current from the cell being measured.

There are many limitations to the use of the hydrogen electrode:

It cannot be used in solutions containing strong oxidants such as ferric iron, dichromates, nitric acid, peroxide or chlorine or reductants, such as sulfurous acid and hydrogen sulfide.

It is affected by the presence of organic compounds which are fairly easily reduced.

It cannot be used successfully in solutions containing cations that fall below hydrogen in the electrochemical series.

Erratic results are obtained in the measurement of unbuffered solutions unless special precautions are taken.

It is troublesome to prepare and maintain.

Since other electrodes more convenient to use now are available, the hydrogen electrode today is used rarely. Nevertheless, it is the ultimate standard for pH measurements.

To avoid some of the difficulties with the hydrogen electrode, the *quinhydrone* electrode was introduced and was popular for a long time, particularly for measurements of acid solutions. The unusual feature of this electrode is that it consists of a piece of gold or platinum wire or foil dipping into the solution to be tested, in which has been dissolved a small quantity of quinhydrone. A calomel electrode may be used for reference, just as in determinations with the hydrogen electrode.

Quinhydrone consists of an equimolecular mixture of quinone and hydroquinone; the relationship between these substances and hydrogen-ion concentration is

$$\text{Quinone} + 2\text{ hydrogen ions} + 2\text{ electrons} \rightleftharpoons \text{hydroquinone}$$

In a solution containing hydrogen ions the potential of the quinhydrone electrode is related logarithmically to hydronium-ion concentration if the ratio of the hydroquinone concentration to that of quinone is constant and practically equal to one. This ratio is maintained in an acid solution containing an excess of quinhydrone, and measurements may be made quickly and accurately; however, quinhydrone cannot be used in solutions more alkaline than pH 8.

An electrode which, because of its simplicity of operation and freedom from contamination or change of the solution being tested, has replaced both the hydrogen and quinhydrone electrodes is the *glass electrode*. It functions by virtue of the fact that when a thin membrane of a special composition of glass separates two solutions of different pH there is developed across the membrane a potential difference which depends on the pH of both solutions. If the pH of one of the solutions is known, the other may be calculated from the potential difference. In practice, the glass electrode usually consists of a bulb of the special glass fused to the end of a tube of ordinary glass. Inside the bulb is placed a solution of known pH, in contact with an internal silver–silver chloride or other electrode. This glass electrode and another reference electrode are immersed in the solution to be tested and the potential difference is measured. A potentiometer providing electronic amplification of the small current produced is employed. The modern instruments available permit reading the pH directly and provide also for compensation of variations due to temperature in the range of 0–50° and to the small but variable asymmetry potential inherent in the glass electrode.

Pharmaceutical Significance

In the broad realm of knowledge concerning the preparation and action of drugs few, if any, variables are so important as pH. For the purpose of this presentation, four principal types of pH-dependence of drug systems will be discussed: solubility, stability, activity and absorption.

Drug Solubility

If a salt, NaA, is added to water to give a concentration C_s, the following reactions occur

$$\text{Na}^+\text{A}^- \xrightarrow{\text{H}_2\text{O}} \text{Na}^+ + \text{A}^-$$

$$\text{A}^- + \text{H}_2\text{O} \rightleftharpoons \text{HA} + \text{OH}^-$$

If the pH of the solution is lowered, more of the A^- would be converted to the unionized acid, HA, in accordance with Le Chatelier's principle. Eventually, a pH will be obtained, below which the amount of HA formed exceeds its aqueous solubility, S_0, and the acid will precipitate from solution; this pH can be designated as pH_p. At this point, at which the amount of HA formed just equals S_0, a mass balance on the total amount of drug in solution yields

$$C_s = [\text{HA}] + [\text{A}^-] = S_0 + [\text{A}^-] \tag{92}$$

Replacing $[A^-]$ as a function of hydronium-ion concentration gives

$$C_s = S_0 + \frac{K_a C_s}{[H_3O^+]_p + K_a} \tag{93}$$

where K_a is the ionization constant for the conjugate acid, HA, and $[H_3O^+]_p$ refers to the hydronium-ion concentration above which precipitation will occur. This equation can be rearranged to give

$$[H_3O^+]_p = K_a \frac{S_0}{C_s - S_0} \tag{94}$$

Taking logarithms gives

$$pH_p = pK_a + \log \frac{C_s - S_0}{S_0} \tag{95}$$

Thus, the pH below which precipitation occurs is a function of the amount of salt added initially, the pK_a and the solubility of the free acid formed from the salt.

The analogous equation for salts of weak bases and strong acids (such as pilocarpine hydrochloride, cocaine hydrochloride or codeine phosphate) would be

$$pH_p = pK_a + \log \frac{S_0}{C_s - S_0} \tag{96}$$

in which pK_a refers to the protonated form of the weak base.

Example—Below what pH will free phenobarbital begin to precipitate from a solution initially containing 1.3 g of sodium phenobarbital/100 mL at 25°? The molar solubility of phenobarbital is 0.0050 and its pK_a is 7.41. The molecular weight of sodium phenobarbital is 254.

The molar concentration of salt initially added is

$$C_s = \frac{g/L}{\text{mol wt}} = \frac{13}{254} = 0.051\ M$$

$$pH_p = 7.41 + \log \frac{0.051 - 0.005}{0.005}$$

$$= 7.41 + 0.96 = 8.37$$

Example—Above what pH will free cocaine begin to precipitate from a solution initially containing 0.0294 mole of cocaine hydrochloride/L? The pK_b of cocaine is 5.59, and its molar solubility is 5.60×10^{-3}.

$$pK_a = pK_w - pK_b = 14.00 - 5.59 = 8.41$$

$$pH_p = 8.41 + \log \frac{0.0056}{0.0294 - 0.0056}$$

$$= 8.41 + (-0.63) = 7.78$$

Drug Stability

One of the most diversified and fruitful areas of study is the investigation of the effect of hydrogen-ion concentration on the stability or, in more general terms, the reactivity of pharmaceutical systems. The evidence for enhanced stability of systems when these are maintained within a narrow range of pH, as well as of progressively decreasing stability as the pH departs from the optimum range, is abundant. Stability (or instability) of a system may result from gain or loss of a proton (hydrogen ion) by a substrate molecule—often accompanied by an electronic rearrangement—which reduces (or increases) the reactivity of the molecule. Instability results when the substance desired to remain unchanged is converted to one or more other, unwanted, substances. In aqueous solution, instability may arise through the catalytic effect of acids or bases, the former by transferring a proton to the substrate molecule, the latter by accepting a proton.

Specific illustrations of the effect of hydrogen-ion concentration on the stability of medicinals are myriad; only a few will be given here, these being chosen to show the importance of pH adjustment of solutions that require sterilization.

Morphine solutions are not decomposed during a 60-min exposure at a temperature of 100° if the pH is less than 5.5; neutral and alkaline solutions, however, are highly unstable. Minimum hydrolytic decomposition of solutions of cocaine occurs in the range of pH of 2 to 5; in one study a solution of cocaine hydrochloride, initially at a pH of 5.7, remained stable during 2 months (although the pH dropped to 4.2 in this time), while another solution buffered to about pH 6 underwent approximately 30% hydrolysis in the same time. Similarly, solutions of procaine hydrochloride containing some hydrochloric acid showed no appreciable decomposition; when dissolved in water alone, 5% of the procaine hydrochloride hydrolyzed, while when buffered to pH 6.5, from 19 to 35% underwent decomposition by hydrolysis. Solutions of thiamine hydrochloride may be sterilized by autoclaving without appreciable decomposition if the pH is below 5; above this, thiamine hydrochloride is unstable.

The stability of many disperse systems, and especially of certain emulsions, is often pH-dependent. Information concerning specific emulsion systems, and the effect of pH upon them, may be found in Chapter 19.

Drug Activity

Drugs that are weak acids or weak bases, and hence may exist in ionized or nonionized form (or a mixture of both), may be active in one form but not in the other; often such drugs have an optimum pH range for maximum activity. Thus, mandelic acid, benzoic acid or salicylic acid have pronounced antibacterial activity in nonionized form but have practically no such activity in ionized form. Accordingly, these substances require an acid environment to function effectively as antibacterial agents. For example, sodium benzoate is effective as a preservative in 4% concentration at pH 7.0, in 0.06 to 0.1% concentration at pH 3.5 to 4.0 and in 0.02 to 0.03% concentration at pH 2.3 to 2.4. Other antibacterial agents, on the other hand, are active principally, if not entirely, in cationic form. Included in this category are the acridines and quaternary ammonium compounds.

Drug Absorption

The degree of ionization and lipoid solubility of a drug are two important factors that determine the rate of absorption of drugs from the gastrointestinal tract and, indeed, their passage through cellular membranes generally. Drugs which are weak organic acids or bases, and which in nonionized form are soluble in lipids, apparently are absorbed through cellular membranes by virtue of the lipoidal nature of the membranes. Completely ionized drugs, on the other hand, are absorbed poorly, if at all. Rates of absorption of a variety of drugs are related to their ionization constants and in many cases may be predicted quantitatively on the basis of this relationship. Thus, not only the degree of the acidic or basic character of a drug but consequently also the pH of the physiological medium (gastric or intestinal fluid, plasma, cerebrospinal fluid, etc) in which a drug is dissolved or dispersed—since this pH determines the extent to which the drug will be converted to ionic or nonionic form—become important parameters of drug absorption. Further information on drug absorption is given in Chapter 35.

References

1. Benet LZ, Goyan JE: *J Pharm Sci 54:* 1179, 1165.
2. Riegelman S *et al: Ibid 51:* 129, 1962.
3. Niebergall PJ *et al: Ibid 61:* 232, 1972.

Bibliography

Freiser H, Fernando Q: *Ionic Equilibria in Analytical Chemistry*, Wiley, New York, 1966.

CHAPTER 19

Disperse Systems

George Zografi, PhD
Professor
School of Pharmacy, University of Wisconsin
Madison, WI 53706

Hans Schott, PhD
Professor of Pharmaceutics and Colloid Chemistry
School of Pharmacy, Temple University
Philadelphia, PA 19140

James Swarbrick, DSc, PhD
Professor and Chairman
Division of Pharmaceutics
School of Pharmacy, University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-7360

Interfacial Phenomena

Very often it is desirable or necessary in the development of pharmaceutical dosage forms to produce multiphasic dispersions by mixing together two or more ingredients which are not mutually miscible and capable of forming homogeneous solutions. Examples of such dispersions include suspensions (solid in liquid), emulsions (liquid in liquid) and foams (vapor in liquids). Because these systems are not homogeneous and thermodynamically stable, over time they will show some tendency to separate on standing to produce the minimum possible surface area of contact between phases. Thus, suspended particles agglomerate and sediment, emulsified droplets cream and coalesce and the bubbles dispersed in foams collapse, to produce unstable and nonuniform dosage forms. In this chapter the fundamental physical chemical properties of dispersed systems will be discussed, along with the principles of interfacial and colloidal physics and chemistry which underly these properties.

Interfacial Forces and Energetics

In the bulk portion of each phase, molecules are attracted to each other equally in all directions, such that no resultant forces are acting on any one molecule. The strength of these forces determines whether a substance exists as a vapor, liquid or solid at a particular temperature and pressure.

At the boundary between phases, however, molecules are acted upon unequally since they are in contact with other molecules exhibiting different forces of attraction. For example, the primary intermolecular forces in water are due to hydrogen bonds, whereas those responsible for intermolecular bonding in hydrocarbon liquids, such as mineral oil, are due to London dispersion forces.

Because of this, molecules situated at the interface contain potential forces of interaction which are not satisfied relative to the situation in each bulk phase. In liquid systems such unbalanced forces can be satisfied by spontaneous movement of molecules from the interface into the bulk phase. This leaves fewer molecules per unit area at the interface (greater intermolecular distance) and reduces the actual contact area between dissimilar molecules.

Any attempt to reverse this process by increasing the area of contact between phases, ie, bringing more molecules into the interface, causes the interface to resist expansion and to behave as though it is under a tension everywhere in a tangential direction. The force of this tension per unit length of interface generally is called the interfacial tension, except when dealing with the air–liquid interface, where the terms surface and surface tension are used.

To illustrate the presence of a tension in the interface, consider an experiment where a circular metal frame, with a looped piece of thread loosely tied to it, is dipped into a liquid. When removed and exposed to the air, a film of liquid will be stretched entirely across the circular frame, as when one uses such a frame to blow soap bubbles. Under these conditions (Fig 19-1*A*), the thread will remain collapsed. If now a heated needle is used to puncture and remove the liquid film from within the loop (Fig 19-1*B*), the loop will stretch spontaneously into a circular shape.

The result of this experiment demonstrates the spontaneous reduction of interfacial contact between air and the liquid remaining and, indeed, that a tension causing the loop to remain extended exists parallel to the interface. The circular shape of the loop indicates that the tension in the plane of the interface exists at right angles or normal to every part of the looped thread. The total force on the entire loop divided by the circumference of the circle, therefore, represents the tension per unit distance of surface, or the surface tension.

Just as work is required to extend a spring under tension, work should be required to reverse the process seen in Figs 19-1*A* and *B*, thus bringing more molecules to the interface. This may be seen quantitatively by considering an experiment where tension and work may be measured directly. Assume that we have a rectangular wire with one movable side (Fig 19-2). Assume further that by dipping this wire into a liquid, a film of liquid will form within the frame when it is removed and exposed to the air. As seen earlier in Fig 19-1, since it comes in contact with air, the liquid surface will tend to contract with a force, F, as molecules leave the surface for the bulk. To keep the movable side in equilibrium, an equal force must be applied to oppose this tension in the surface. We then may define the surface tension, γ, of the liquid as $F/2l$, where $2l$ is the distance of surface over which F is operating ($2l$ since there are two surfaces, top and bottom). If the surface is expanded by a very small distance, Δx, one can then estimate that the work done is

$$W = F\Delta x \quad (1)$$

and therefore

$$W = \gamma 2l\Delta x \quad (2)$$

Dr Zografi authored the section on *Interfacial Phenomena*. *Dr Schott* authored the section on Colloidal Dispersions. *Dr Swarbrick* authored the section on Particle Phenomena and Coarse Dispersions.

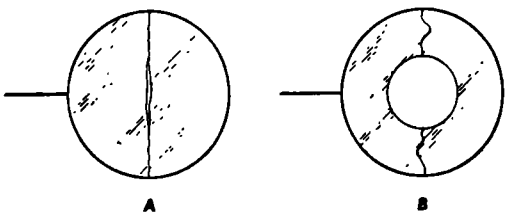


Fig 19-1. A circular wire frame with a loop of thread loosely tied to it: (*A*) a liquid film on the wire frame with a loop in it; (*B*) the film inside the loop is broken.[1]

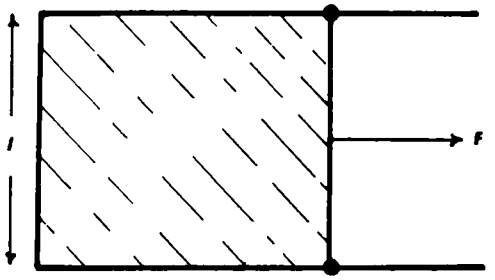


Fig 19-2. A movable wire frame containing a film of liquid being expanded with a force, *F*.

Since

$$\Delta A = 2l\Delta x \tag{3}$$

where ΔA is the change in area due to the expansion of the surface, we may conclude that

$$W = \gamma \Delta A \tag{4}$$

Thus, the work required to create a unit area of surface, known as the surface free energy/unit area, is equivalent to the surface tension of a liquid system, and the greater the area of interfacial contact between phases, the greater the free-energy increase for the total system. Since a prime requisite for equilibrium is that the free energy of a system be at a minimum, it is not surprising to observe that phases in contact tend to reduce area of contact spontaneously.

Liquids, being mobile, may assume spherical shapes (smallest interfacial area for a given volume), as when ejected from an orifice into air or when dispersed into another immiscible liquid. If a large number of drops are formed, further reduction in area can occur by having the drops coalesce, as when a foam collapses or when the liquid phases making up an emulsion separate.

Surface tension is expressed in units of dynes/cm, while surface free energy is expressed in ergs/cm^2. Since an erg is a dyne-cm, both sets of units are equivalent.

Values for the surface tension of a variety of liquids are given in Table I, while interfacial tension values for various liquids against water are given in Table II. Other combinations of immiscible phases could be given but most heterogeneous systems encountered in pharmacy usually contain water. Values for these tensions are expressed for a particular temperature. Since an increased temperature increases the thermal energy of molecules, the work required to bring molecules to the interface should be less, and thus the surface and interfacial tension will be reduced. For example, the surface tension of water at 0° is 76.5 dynes/cm and 63.5 dynes/cm at 75°.

As would be expected from the discussion so far, the relative values for surface tension should reflect the nature of intermolecular forces present; hence, the relatively large values for mercury (metallic bonds) and water (hydrogen bonds), and the lower values for benzene, chloroform, carbon tetrachloride and the *n*-alkanes. Benzene with π electrons

Table I—Surface Tension of Various Liquids at 20°

Substance	Surface tension, dynes/cm
Mercury	476
Water	72.8
Glycerin	63.4
Oleic acid	32.5
Benzene	28.9
Chloroform	27.1
Carbon tetrachloride	26.8
1-Octanol	26.5
Hexadecane	27.4
Dodecane	25.4
Decane	23.9
Octane	21.8
Heptane	19.7
Hexane	18.0
Perfluoroheptane	11.0
Nitrogen (at 75°K)	9.4

Table II—Interfacial Tension of Various Liquids against Water at 20°

Substance	Interfacial tension, dynes/cm
Decane	52.3
Octane	51.7
Hexane	50.8
Carbon tetrachloride	45.0
Chloroform	32.8
Benzene	35.0
Mercury	428
Oleic acid	15.6
1-Octanol	8.51

exhibits a higher surface tension than the alkanes of comparable molecular weight, but increasing the molecular weight of the alkanes (and hence intermolecular attraction) increases their surface tension closer to that of benzene. The lower values for the more nonpolar substances, perfluoroheptane and liquid nitrogen, demonstrate this point even more strongly.

Values of interfacial tension should reflect the differences in chemical structure of the two phases involved; the greater the tendency to interact, the less the interfacial tension. The 20-dynes/cm difference between air–water tension and that at the octane–water interface reflects the small but significant interaction between octane molecules and water molecules at the interface. This is seen also in Table II, by comparing values for octane and octanol, oleic acid and the alkanes, or chloroform and carbon tetrachloride.

In each case the presence of chemical groups capable of hydrogen bonding with water markedly reduces the interfacial tension, presumably by satisfying the unbalanced forces at the interface. These observations strongly suggest that molecules at an interface arrange themselves or orient so as to minimize differences between bulk phases.

That this occurs even at the air–liquid interface is seen when one notes the relatively low surface-tension values of very different chemical structures such as the *n*-alkanes, octanol, oleic acid, benzene and chloroform. Presumably, in each case, the similar nonpolar groups are oriented toward the air with any polar groups oriented away toward the bulk phase. This tendency for molecules to orient at an interface is a basic factor in interfacial phenomena and will be discussed more fully in succeeding sections.

Solid substances such as metals, metal oxides, silicates and salts, all containing polar groups exposed at their surface, may be classified as high-energy solids, whereas nonpo-

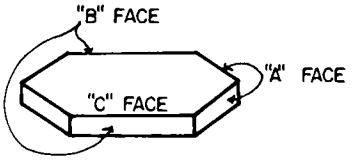


Fig 19-3. Adipic acid crystal showing various faces.[2]

Table III—Values of γ_{sv} for Solids of Varying Polarity

Solid	γ_{sv} (dynes/cm)
Teflon	19.0
Paraffin	25.5
Polyethylene	37.6
Polymethyl methacrylate	45.4
Nylon	50.8
Indomethacin	61.8
Griseofulvin	62.2
Hydrocortisone	68.7
Sodium Chloride	155
Copper	1300

lar solids such as carbon, sulfur, glyceryl tristearate, polyethylene and polytetrafluoroethylene (Teflon) may be classified as low-energy solids. It is of interest to measure the surface free energy of solids; however, the lack of mobility of molecules at the surface of solids prevents the observation and direct measurement of a surface tension. It is possible to measure the work required to create new solid surface by cleaving a crystal and measuring the work involved. However, this work not only represents free energy due to exposed groups but also takes into account the mechanical energy associated with the crystal (ie, plastic and elastic deformation and strain energies due to crystal structure and imperfections in that structure).

Also contributing to the complexity of a solid surface is the heterogeneous behavior due to the exposure of different crystal faces, each having a different surface free energy/unit area. For example, adipic acid, $HOOC(CH_2)_4COOH$, crystallizes from water as thin hexagonal plates with three different faces, as shown in Fig 19-3. Each unit cell of such a crystal contains adipic acid molecules oriented such that the hexagonal planes (faces) contain exposed carboxyl groups, while the sides and edges (A and B faces) represent the side view of the carboxyl and alkyl groups, and thus are quite nonpolar. Indeed, interactions involving these different faces reflect the differing surface free energies.[2]

Other complexities associated with solid surfaces include surface roughness, porosity and the defects and contamination produced during a recrystallization or comminution of the solid. In view of all these complications, surface free energy values for solids, when reported, should be regarded as average values, often dependent on the method used and not necessarily the same for other samples of the same substance.

In Table III are listed some approximate average values of γ_{sv} for a variety of solids, ranging in polarity from Teflon to copper, obtained by various indirect techniques.

Adhesional and Cohesional Forces

Of prime importance to those dealing with heterogeneous systems is the question of how two phases will behave when brought in contact with each other. It is well known, for instance, that some liquids, when placed in contact with other liquid or solid surfaces, will remain retracted in the form of a drop (known as a lens), while other liquids may exhibit a tendency to spread and cover the surface of this liquid or solid.

Based upon concepts developed to this point, it is apparent that the individual phases will exhibit a tendency to minimize the area of contact with other phases, thus leading to phase separation. On the other hand, the tendency for interaction between molecules at the new interface will offset this to some extent and give rise to the spontaneous spreading of one substance over the other.

In essence, therefore, phase affinity is increased as the forces of attraction between different phases (adhesional forces) become greater than the forces of attraction between molecules of the same phase (cohesional forces). If these adhesional forces become great enough, miscibility will occur and the interface will disappear. The present discussion is concerned only with systems of limited phase affinity, where an interface still exists.

A convenient approach used to express these forces quantitatively involves the use of the terms work of adhesion and work of cohesion.

The work of adhesion, W_a, is defined as the energy per cm^2 required to separate two phases at their boundary and is equal but opposite in sign to the free energy/cm^2 released when the interface is formed. In an analogous manner the work of cohesion for a pure substance, W_c, is the work/cm^2 required to produce two new surfaces, as when separating different phases, but now both surfaces contain the same molecules. This is equal and opposite in sign to the free energy/cm^2 released when the same two pure liquid surfaces are brought together and eliminated.

By convention, when the work of adhesion between two substances, A and B, exceeds the work of cohesion for one substance, eg, B, spontaneous spreading of B over the surface of A should occur with a net loss of free energy equal to the difference between W_a and W_c. If W_c exceeds W_a, no spontaneous spreading of B over A can occur. The difference between W_a and W_c is known as the spreading coefficient, S; only when S is positive will spreading occur.

The values for W_a and W_c (and hence S) may be expressed in terms of surface and interfacial tensions, when one considers that upon separation of two phases, A and B, γ_{AB} ergs of interfacial free energy/cm^2 (interfacial tension) are lost, but that γ_A and γ_B ergs/cm^2 of energy (surface tensions of A and B) are gained; upon separation of bulk phase molecules in an analogous manner, $2\gamma_A$ or $2\gamma_B$ ergs/cm^2 will be gained. Thus

$$W_a = \gamma_A + \gamma_B - \gamma_{AB} \tag{5}$$

and

$$W_c = 2\gamma_A \text{ or } 2\gamma_B \tag{6}$$

For B spreading on the surface of A, therefore

$$S_B = \gamma_A + \gamma_B - \gamma_{AB} - 2\gamma_B \tag{7}$$

or

$$S_B = \gamma_A - (\gamma_B + \gamma_{AB}) \tag{8}$$

Utilizing Eq 8 and values of surface and interfacial tension given in Tables I and II, S can be calculated for three representative substances—decane, benzene, and oleic acid—on water at 20°.

Decane:	$S = 72.8 - (23.9 + 52.3)$	$= -3.4$
Benzene:	$S = 72.8 - (28.9 + 35.0)$	$= 8.9$
Oleic acid:	$S = 72.8 - (32.5 + 15.6)$	$= 24.7$

As expected, relatively nonpolar substances such as decane exhibit negative values of S, whereas the more polar materials yield positive values; the greater the polarity of the mole-

cule, the more positive the value of S. The importance of the cohesive energy of the spreading liquid may be noted also by comparing the spreading coefficients for hexane on water and water on hexane:

$$S_{H/W} = 72.8 - (18.0 + 50.8) = \quad 4.0$$

$$S_{W/H} = 18.0 - (72.8 + 50.8) = -105.6$$

Here, despite the fact that both liquids are the same, the high cohesion and air–liquid tension of water prevents spreading on the low-energy hexane surface, while the very low value for hexane allows spreading on the water surface. This also is seen when comparing the positive spreading coefficient of hexane to the negative value for decane on water.

To see whether spreading does or does not occur, a powder such as talc or charcoal can be sprinkled over the surface of water such that it floats; then, a drop of each liquid is placed on this surface. As predicted, decane will remain as an intact drop, while hexane, benzene and oleic acid will spread out, as shown by the rapid movement of solid particles away from the point where the liquid drop was placed originally.

An apparent contradiction to these observations may be noted for hexane, benzene and oleic acid when more of each substance is added, in that lenses now appear to form even though initial spreading occurred. Thus, in effect a substance does not appear to spread over itself.

It is now established that the spreading substance forms a monomolecular film which creates a new surface having a lower surface free energy than pure water. This arises because of the apparent orientation of the molecules in such a film so that their most hydrophobic portion is oriented towards the spreading phase. It is the lack of affinity between this exposed portion of the spread molecules and the polar portion of the remaining molecules which prevents further spreading.

This may be seen by calculating a final spreading coefficient where the new surface tension of water plus monomolecular film is used. For example, the presence of benzene reduces the surface tension of water to 62.2 dynes/cm so that the final spreading coefficient, S_F, is

$$S_F = 62.2 - (28.9 + 35.0) = -1.7$$

The lack of spreading exhibited by oleic acid should be reflected in an even more negative final spreading coefficient, since the very polar carboxyl groups should have very little affinity for the exposed alkyl chain of the oleic acid film. Spreading so as to form a second layer with polar groups exposed to the air would also seem very unlikely, thus leading to the formation of a lens.

Wetting Phenomena

In the experiment described above it was shown that talc or charcoal sprinkled onto the surface of water float despite the fact that their densities are much greater than that of water. In order for immersion of the solid to occur, the liquid must displace air and spread over the surface of the solid; when liquids cannot spread over a solid surface spontaneously, and, therefore, S, the spreading coefficient, is negative, we say that the solid is not wetted.

An important parameter which reflects the degree of wetting is the angle which the liquid makes with the solid surface at the point of contact (Fig 19-4). By convention, when wetting is complete, the contact angle is zero; in nonwetting situations it theoretically can increase to a value of 180°, where a spherical droplet makes contact with solid at only one point.

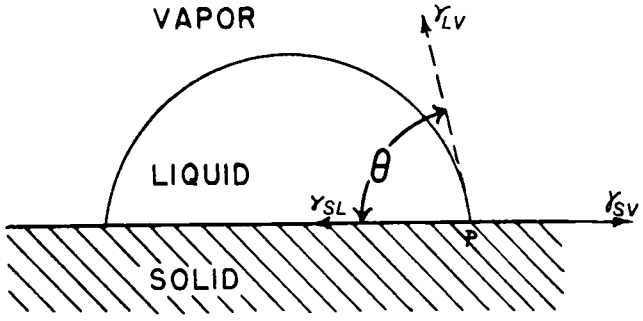


Fig 19-4. Forces acting on a nonwetting liquid drop exhibiting a contact angle of θ.[3]

In order to express contact angle in terms of solid–liquid–air equilibria, one can balance forces parallel to the solid surface at the point of contact between all three phases (Fig 19-4), as expressed in

$$\gamma_{SV} = \gamma_{SL} + \gamma_{LV} \cos \theta \qquad (9)$$

where γ_{SV}, γ_{SL}, and γ_{LV} represent the surface free energy/unit area of the solid–air, solid–liquid, and liquid–air interfaces, respectively. Although difficult to use quantitatively because of uncertainties with γ_{SV} and γ_{SL} measurements, conceptually the equation, known as the Young equation, is useful because it shows that the loss of free energy due to elimination of the air–solid interface by wetting is offset by the increased solid–liquid and liquid–air area of contact as the drop spreads out.

The $\gamma_{LV} \cos \theta$ term arises as the horizontal vectorial component of the force acting along the surface of the drop, as represented by γ_{LV}. Factors tending to reduce γ_{LV} and γ_{SL}, therefore, will favor wetting, while the greater the value of γ_{SV} the greater the chance for wetting to occur. This is seen in Table IV for the wetting of a low-energy surface, paraffin (hydrocarbon), and a higher energy surface, nylon, (polyhexamethylene adipamide). Here, the lower the surface tension of a liquid, the smaller the contact angle on a given solid, and the more polar the solid, the smaller the contact angle with the same liquid.

With Eq 9 in mind and looking at Fig 19-5, it is now possible to understand how the forces acting at the solid-

Table IV—Contact Angle on Paraffin and Nylon for Various Liquids of Differing Surface Tension

Substance	Surface tension, dynes/cm	Contact angle Paraffin	Nylon
Water	72.8	105°	70°
Glycerin	63.4	96°	60°
Formamide	58.2	91°	50°
Methylene iodide	50.8	66°	41°
α-Bromonaphthalene	44.6	47°	16°
tert-Butylnaphthalene	33.7	38°	spreads
Benzene	28.9	24°	"
Dodecane	25.4	17°	"
Decane	23.9	7°	"
Nonane	22.9	spreads	"

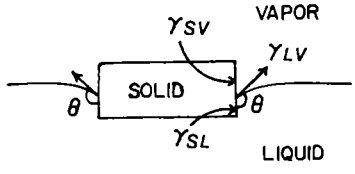


Fig 19-5. Forces acting on a nonwettable solid at the air+liquid+solid interface: contact angle θ greater than 90°.

Table V—Critical Surface Tensions of Various Polymeric Solids

Polymeric Solid	γ_c, Dynes/cm at 20°
Polymethacrylic ester of ϕ'-octanol	10.6
Polyhexafluoropropylene	16.2
Polytetrafluoroethylene	18.5
Polytrifluoroethylene	22
Poly(vinylidene fluoride)	25
Poly(vinyl fluoride)	28
Polyethylene	31
Polytrifluorochloroethylene	31
Polystyrene	33
Poly(vinyl alcohol)	37
Poly(methyl methacrylate)	39
Poly(vinyl chloride)	39
Poly(vinylidene chloride)	40
Poly(ethylene terephthalate)	43
Poly(hexamethylene adipamide)	46

liquid-air interface can cause a dense nonwetted solid to float if γ_{SL} and γ_{LV} are large enough relative to γ_{SV}.

The significance of reducing γ_{LV} was first developed empirically by Zisman when he plotted cos θ vs the surface tension of a series of liquids and found that a linear relationship, dependent on the solid, was obtained. When such plots are extrapolated to cos θ equal to one or a zero contact angle, a value of surface tension required to just cause complete wetting is obtained. Doing this for a number of solids, it was shown that this surface tension (known as the critical surface tension, γ_c) parallels expected solid surface energy γ_{SV}; the lower γ_c, the more nonpolar the surface.

Table V indicates some of these γ_c values for different surface groups, indicating such a trend. Thus, water with a surface tension of about 72 dynes/cm will not wet polyethylene (γ_c = 31 dynes/cm), but heptane with a surface tension of about 20 dynes/cm will. Likewise, Teflon (polytetrafluoroethylene) (γ_c = 19) is not wetted by heptane but is wetted by perfluoroheptane with a surface tension of 11 dynes/cm.

One complication associated with the wetting of high-energy surfaces is the lack of wetting after the initial formation of a monomolecular film by the spreading substance. As in the case of oleic acid spreading on the surface of water, the remaining liquid retracts because of the low-energy surface produced by the oriented film. This phenomenon, often called autophobic behavior, is an important factor in many systems of pharmaceutical interest since many solids, expected to be wetted easily by water, may be rendered hydrophobic if other molecules dissolved in the water can form these monomolecular films at the solid surface.

Capillarity

Because water shows a strong tendency to spread out over a polar surface such as clean glass (contact angle 0°), one would expect to observe the meniscus which forms when water is contained in a glass vessel such as a pipet or buret. This behavior is accentuated dramatically if a fine-bore capillary tube is placed into the liquid (Fig 19-6); not only will the wetting of the glass produce a more highly curved meniscus, but the level of the liquid in the tube will be appreciably higher than the level of the water in the beaker.

The spontaneous movement of a liquid into a capillary or narrow tube due to surface forces is defined as capillarity and is responsible for a number of important processes involving the penetration of liquids into porous solids. In contrast to water in contact with glass, if the same capillary is placed into mercury (contact angle on glass: 130°), not

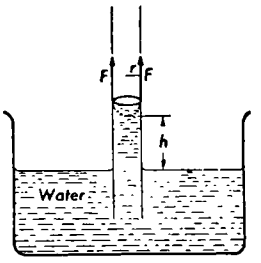


Fig 19-6. Capillary rise for a liquid exhibiting zero contact angle [1]

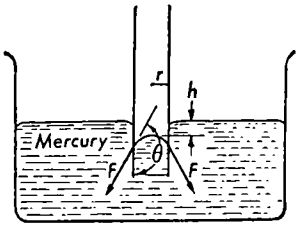


Fig 19-7. Capillary fall for a liquid exhibiting a contact angle, θ, which is greater than 90°.[1]

only will the meniscus be inverted (see Fig 19-7), but the level of the mercury in the capillary will be lower than in the beaker. In this case one does not expect mercury or other *nonwetting* liquids to easily penetrate pores unless external forces are applied.

To quantitate the factors giving rise to the phenomenon of capillarity, let us consider the case of a liquid which rises to a height, h, above the bulk liquid in a capillary having a radius, r. If (as shown in Fig 19-6) the contact angle of water on glass is zero, a force, F, will act upward and vertically along the circle of liquid–glass contact. Based upon the definition of surface tension this force will be equal to the surface tension, γ, multiplied by the circumference of the circle, $2\pi r$. Thus

$$F = \gamma 2\pi r \tag{10}$$

This force upward must support the column of water, and since the mass, m, of the column is equal to the density, d, multiplied by the volume of the column, $\pi r^2 h$, the force W opposing the movement upward will be

$$W = mg = \pi r^2 dgh \tag{11}$$

where g is the gravity constant.

Equating the two forces at equilibrium gives

$$\pi r^2 dgh = \gamma 2\pi r \tag{12}$$

so that

$$h = \frac{2\gamma}{rdg} \tag{13}$$

Thus, the greater the surface tension and the finer the capillary radius, the greater the rise of liquid in the capillary.

If the contact angle of liquid is not zero (as shown in Fig 19-8), the same relationship may be developed, except the

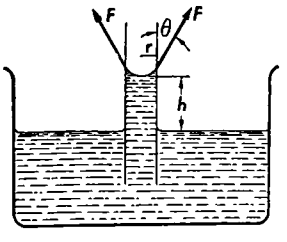


Fig 19-8. Capillary rise for a liquid exhibiting a contact angle, θ, which is greater than zero but less than 90°.[1]

vertical component of F which opposes the weight of the column is $F \cos \theta$ and, therefore

$$h = \frac{2\gamma \cos \theta}{rdg} \quad (14)$$

This indicates the very important fact that if θ is less than 90°, but greater than 0°, the value of h will decrease with increasing contact angle until at 90° ($\cos \theta = 0$), $h = 0$. Above 90°, values of h will be negative, as indicated in Fig 19-7 for mercury. Thus, based on these equations we may conclude that capillarity will occur spontaneously in a cylindrical pore even if the contact angle is greater than zero, but it will not occur at all if the contact angle becomes 90° or more. In solids with irregularly shaped pores the relationships between parameters in Eq 14 will be the same, but they will be more difficult to quantitate because of nonuniform changes in pore radius throughout the porous structure.

Pressure Differences across Curved Surfaces

From the preceding discussion of capillarity another important concept follows. In order for the liquid in a capillary to rise spontaneously it must develop a higher pressure than the lower level of the liquid in the beaker. However, since the system is open to the atmosphere, both surfaces are in equilibrium with the atmospheric pressure. In order to be raised above the level of liquid in the beaker and produce a hydrostatic pressure equal to hgd, the pressure just below the liquid meniscus, in the capillary, P_1, must be less than that just below the flat liquid surface, P_0, by hgd, and therefore

$$P_0 - P_1 = hgd \quad (15)$$

Since, according to Eq 14

$$h = \frac{2\gamma \cos \theta}{rgd}$$

then

$$P_0 - P_1 = \frac{2\gamma \cos \theta}{r} \quad (16)$$

For a contact angle of zero, where the radius of the capillary is the radius of the hemisphere making up the meniscus,

$$P_0 - P_1 = \frac{2\gamma}{r} \quad (17)$$

The consequences of this relationship (known as the Laplace equation) are important for any curved surface when r becomes very small and γ is relatively significant. For example, a spherical droplet of air formed in a bulk liquid and having a radius, r, will have a greater pressure on the inner concave surface than on the convex side, as expressed in Eq 17.

Another direct consequence of what Eq 17 expresses is the fact that very small droplets of liquid, having highly curved surfaces, will exhibit a higher vapor pressure, P, than that observed over a flat surface of the same liquid at P'. The equation (Eq. 18) expressing the ratio of P/P' to droplet radius, r, and surface tension, γ, is called the Kelvin equation where

$$\log P/P' = \frac{2\gamma M}{2.303RT\rho r} \quad (18)$$

and M is the molecular weight, R the gas constant in ergs per mole per degree, T is temperature and ρ is the density in g/cm^3. Values for the ratio of vapor pressures are given in Table VI for water droplets of varying size. Such ratios indicate why it is possible for very fine water droplets in clouds to remain uncondensed despite their close proximity to one another.

Table VI—Ratio of Observed Vapor Pressure to Expected Vapor Pressure of Water at 25° with Varying Droplet Size

P/P' [a]	Droplet size, μm
1.001	1
1.01	0.1
1.1	0.01
2.0	0.005
3.0	0.001
4.2	0.00065
5.2	0.00060

[a] P is the observed vapor pressure and P' is the expected value for "bulk" water.

This same behavior may be seen when measuring the solubility of very fine solid particles since both vapor pressure and solubility are measures of the escaping tendency of molecules from a surface. Indeed, the equilibrium solubility of extremely small particles has been shown to be greater than the usual value noted for coarser particles; the greater the surface energy and smaller the particles, the greater this effect.

Adsorption

Vapor Adsorption on Solid Surfaces

It was suggested earlier that a high surface or interfacial free energy may exist at a solid surface if the unbalanced forces at the surface and the area of exposed groups are quite great.

Substances such as metals, metal oxides, silicates, and salts—all containing exposed polar groups—may be classified as high-energy or hydrophilic solids; nonpolar solids such as carbon, sulfur, polyethylene, or Teflon (polytetrafluoroethylene) may be classified as low-energy or hydrophobic solids (Table III). Whereas liquids satisfy their unbalanced surface forces by changes in shape, pure solids (which exhibit negligible surface mobility) must rely on reaction with molecules either in the vapor state or in a solution which comes in contact with the solid surface to accomplish this.

Vapor adsorption is the simplest model demonstrating how solids reduce their surface free energy in this manner.

Depending on the chemical nature of the adsorbent (solid) and the adsorbate (vapor), the strength of interaction between the two species may vary from strong specific chemical bonding to interactions produced by the weaker more nonspecific London dispersion forces. Ordinarily, these latter forces are those responsible for the condensation of relatively nonpolar substances such as N_2, O_2, CO_2 or hydrocarbons.

When chemical reaction occurs, the process is called chemisorption; when dispersion forces predominate, the term physisorption is used. Physisorption occurs at temperatures approaching the liquefaction temperature of the vapor, whereas, for chemisorption, temperatures depend on the particular reaction involved. Water-vapor adsorption to various polar solids can occur at room temperature through hydrogen-bonding, with binding energies intermediate to physisorption and chemisorption.

In order to study the adsorption of vapors onto solid surfaces one must measure the amount of gas adsorbed/unit area or unit mass of solid, at different pressures of gas. Since such studies usually are conducted at constant temperature, plots of volume adsorbed vs pressure are referred to as adsorption isotherms. If the physical or chemical adsorption process is monomolecular, the adsorption iso-

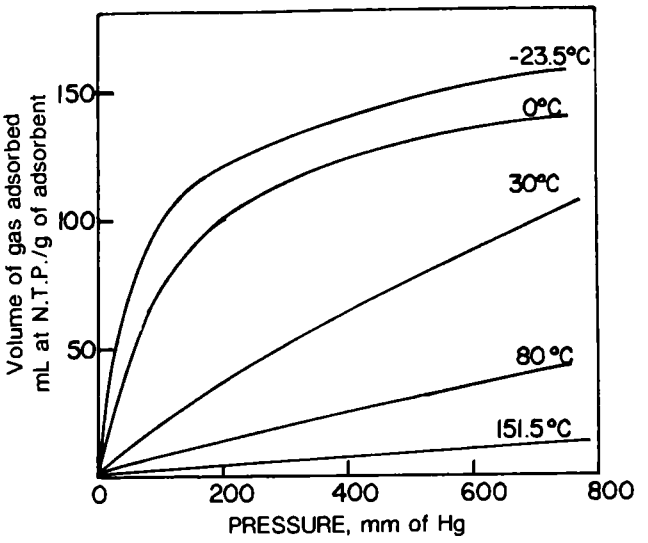


Fig 19-9 Adsorption isotherms for ammonia on charcoal.[4]

therm should look like those shown in Fig 19-9. Note the significant increase in adsorption with increasing pressure, followed by a leveling off. This leveling off is due either to a saturation of available specific chemical groups, as in chemisorption, or to the entire available surface being covered by physically adsorbed molecules. Note also the reduction in adsorption with increasing temperature which occurs because the adsorption process is exothermic. Often in the case of physical adsorption at low temperatures, after adsorption levels off, a marked increase in adsorption occurs, presumably due to multilayered adsorption. In this case vapor molecules essentially condense upon themselves as the liquefaction pressure of the vapor is approached. Figure 19-10 illustrates one type of isotherm generally seen with multilayered physisorption.

In order to have some quantitative understanding of the adsorption process and to be able to compare different systems, two factors must be evaluated; it is important to know what the capacity of the solid is or what the maximum amount of adsorption is under a given set of conditions and what the affinity of a given substance is for the solid surface or how readily does it adsorb for a given amount of pressure? In effect, this second term is the equilibrium constant for the process.

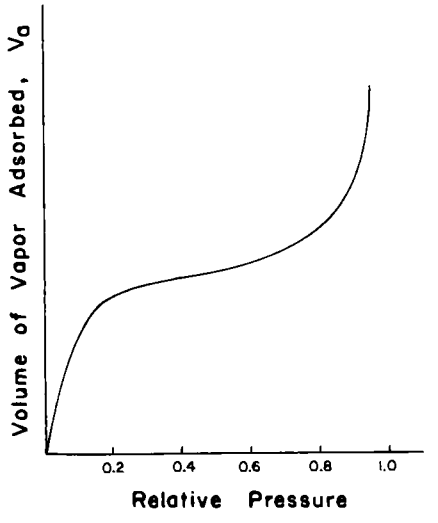


Fig 19-10. Typical plot for multilayer physical adsorption of a vapor on a solid surface.

A significant development along these lines was introduced by Langmuir when he proposed his theory of monomolecular adsorption. He postulated that for adsorption to occur a solid must contain uniform adsorption sites, each capable of holding a gas molecule. Molecules colliding with the surface may bounce off elastically or they may remain in contact for a period of time. It is this contact over a period of time that Langmuir termed adsorption.

Two major assumptions were made in deriving the equation: (1) only those molecules striking an empty site can be adsorbed, hence, only monomolecular adsorption occurs, and (2) the forces of interaction between adsorbed molecules are negligible and, therefore, the probability of a molecule adsorbing onto or desorbing from any site is independent of the surrounding sites.

The derivation of the equation is based upon the relationship between the rate of adsorption and desorption, since at equilibrium the two rates must be equal. Let μ equal the number of molecules striking each sq cm of surface/sec. From the kinetic theory of gases

$$\mu = \frac{p}{(2\pi mkT)^{1/2}} \tag{19}$$

where p is the gas pressure, m is the mass of the molecule, k is the Boltzmann gas constant, and T is the absolute temperature. Thus, the greater p, the greater the number of collisions. Let α equal the fraction of molecules which will be held by the surface; then $\alpha\mu$ is equal to the rate of adsorption on the bare surface. However, if θ is the fraction of the surface already covered, the rate of adsorption actually will be

$$R_a = \alpha\mu(1 - \theta) \tag{20}$$

In a similar manner the rate of molecules leaving the surface can be expressed as

$$R_d = \gamma\theta \tag{21}$$

where γ is the rate at which molecules can leave the surface and θ represents the number of molecules available to desorb. The value of γ strongly depends on the energy associated with adsorption; the greater the binding energy, the lower the value of γ. At equilibrium, $R_a = R_d$ and

$$\gamma\theta = \alpha\mu(1 - \theta) \tag{22}$$

Isolating the variable term, p, and combining all constants into k, the equation can be written as

$$\theta = \frac{kp}{1 + kp} \tag{23}$$

and, since θ may be expressed as

$$\theta = \frac{V_a}{V_m} \tag{24}$$

where V_a is the volume of gas adsorbed and V_m is the volume of gas covering all of the sites, Eq. 23 may be written as

$$V_a = \frac{V_m kp}{1 + kp} \tag{25}$$

A test of fit to this equation can be made by expressing it in linear form

$$\frac{p}{V_a} = \frac{1}{V_m k} + \frac{p}{V_m} \tag{26}$$

The value of k is, in effect, the equilibrium constant and may be used to compare affinities of different substances for the solid surface. The value of V_m is valuable since it indicates the maximum number of sites available for adsorption. In the case of physisorption the maximum number of sites is

actually the total surface area of the solid and, therefore, the value of V_m can be used to estimate surface area if the volume and area/molecule of vapor are known.

Since physisorption most often involves some multilayered adsorption, an equation, based on the Langmuir equation, the B.E.T. equation, is normally used to determine V_m and solid surface areas. Equation 27 is the B.E.T. equation:

$$V_a = \frac{V_m c p}{(p_0 - p)[1 + (C - 1)(p/p_0]} \qquad (27)$$

where c is a constant and p_0 is the vapor pressure of the adsorbing substance.[5] The most widely used vapor for this purpose is nitrogen, which adsorbs nonspecifically on most solids near its boiling point at −195° and appears to occupy about 16 Å^2/molecule on a solid surface.

Adsorption from Solution

By far one of the most important aspects of interfacial phenomena encountered in pharmaceutical systems is the tendency for substances dissolved in a liquid to adsorb to various interfaces. Adsorption from solution is generally more complex than that from the vapor state because of the influence of the solvent and any other solutes dissolved in the solvent. Although such adsorption is generally limited to one molecular layer, the presence of other molecules often makes the interpretation of adsorption mechanisms much more difficult than for chemisorption or physisorption of a vapor. Since monomolecular adsorption from solution is so widespread at all interfaces, we will first discuss the nature of monomolecular films and then return to a discussion of adsorption from solution.

Insoluble Monomolecular Films

It was suggested above that molecules exhibiting a tendency to spread out at an interface might be expected to orient so as to reduce the interfacial free energy produced by the presence of the interface. Direct evidence for molecular orientation has been obtained from studies dealing with the spreading on water of insoluble polar substances containing long hydrocarbon chains, eg, fatty acids.

In the late 19th century Pockels and Rayleigh showed that a very small amount of olive or castor oil—when placed on the surface of water—spreads out, as discussed above. If the amount of material was less than could physically cover the entire surface only a slight reduction in the surface tension of water was noted. However, if the surface was compressed between barriers, as shown in Fig 19-11, the surface tension was reduced considerably.

Devaux extended the use of this technique by dissolving small amounts of solid in volatile solvents and dropping the solution onto a water surface. After assisting the water-insoluble molecules to spread, the solvent evaporated, leaving a surface film containing a known amount of solute.

Compression and measurement of surface tension indicated that a maximum reduction of surface was reached when the number of molecules/unit area was reduced to a value corresponding to complete coverage of the surface. This suggested that a monomolecular film forms and that surface tension is reduced upon compression because contact between air and water is reduced by the presence of the film molecules. Beyond the point of closest packing the film apparently collapses very much as a layer of corks floating on water would be disrupted when laterally compressed beyond the point of initial physical contact.

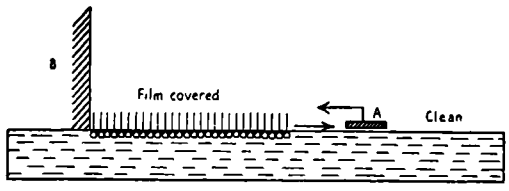


Fig 19-11. Insoluble monomolecular film compressed between a fixed barrier, *B*, and a movable barrier, *A*.[6]

Using a refined quantitative technique based on these studies, Langmuir[7] spread films of pure fatty acids, alcohols, and esters on the surface of water. Comparing a series of saturated fatty acids, differing only in chain length, he found that the area/molecule at collapse was independent of chain length, corresponding to the cross-sectional area of a molecule oriented in a vertical position (see Fig 19-11). He further concluded that this molecular orientation involved association of the polar carboxyl group with the water phase and the nonpolar acyl chain out towards the vapor phase.

In addition to the evidence for molecular orientation, Langmuir's work with surface films revealed that each substance exhibits film properties which reflect the interactions between molecules in the surface film. This is best seen by plotting the difference in surface tension of the clean surface, γ_0, and that of the surface covered with the film, γ, vs the area/molecule, A, produced by film compression (total area ÷ the number of molecules). The difference in surface tension is called the surface pressure, π, and thus

$$\pi = \gamma_0 - \gamma. \qquad (28)$$

Figure 19-12 depicts such a plot for a typical fatty acid monomolecular film. At areas greater than 50 Å^2/molecule the molecules are far apart and do not cover enough surface to reduce the surface tension of the clean surface to any extent and thus the lack of appreciable surface pressure. Since the molecules in the film are quite free to move laterally in the surface, they are said to be in a two-dimensional "gaseous" or "vapor" state.

As the intermolecular distance is reduced upon compression, the surface pressure rises because the air–water surface is being covered to a greater extent. The rate of change in π with A, however, will depend on the extent of interaction between film molecules; the greater the rate of change, the more "condensed" the state of the film.

In Fig 19-12, from 50 Å^2 to 30 Å^2/molecule, the curve shows a steady increase in π, representative of a two-dimensional "liquid" film, where the molecules become more restricted in their freedom of movement because of interactions. Below 30 Å^2/molecule the increase in π occurs over a narrow range of A, characteristic of closest packing and a two-dimensional "solid" film.

Any factor tending to increase polarity or bulkiness of the molecule—such as increased charge, number of polar groups, reduction in chain length, or the introduction of

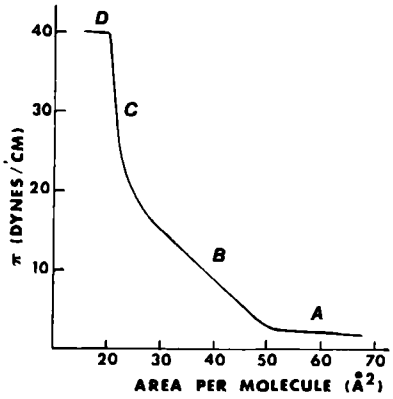


Fig 19-12. A surface pressure–area curve for an insoluble monomolecular film: Region *A*, "gaseous" film; Region *B*, "liquid" film; Region *C*, "solid" film; Region *D*, film collapse

aromatic rings, side chains, and double bonds—should reduce molecular interactions, while the longer the alkyl chain and the less bulky the polar group, the closer the molecules can approach and the stronger the extent of interaction in the film.

Soluble Films and Adsorption from Solution

If a fatty acid exhibits highly "gaseous" film behavior on an aqueous surface, we should expect a relatively small change in π with A over a considerable range of compression. Indeed, for short-chain compounds—eg, lauric acid (12 carbons) and decanoic acid—not only is the change in π small with decreasing A but at a point just before the expected closest packing area the surface pressure becomes constant without any collapse.

If lauric acid is converted to the laurate ion, or if a shorter chain acid such as octanoic acid is used, spreading on water and compression of the surface produces no increase in π; the more polar the molecule (hence, the more "gaseous" the film), the higher the area/molecule where a constant surface pressure occurs.

This behavior may be explained by assuming that polar molecules form monomolecular films when spread on water but that, upon compression, they are caused to enter the aqueous bulk solution rather than to remain as an intact insoluble film. The constant surface pressure with increased compression arises because a constant number of molecules/unit area remain at the surface in equilibrium with dissolved molecules. The extent of such behavior will be greater for substances exhibiting weaker intermolecular interaction and greater water solubility.

Starting from the other direction, it can be shown that short-chain acids and alcohols (when dissolved in water) reduce the surface tension of water, thus producing a surface pressure, just as with insoluble films (see Eq 28). That dissolved molecules are accumulating at the interface in the form of a monomolecular film is suggested from the similarity in behavior to systems where slightly soluble molecules are spread on the surface. For example, compressing the surface of a solution containing "surface-active" molecules has no effect on the initial surface pressure, whereas increasing bulk-solution concentration tends to increase surface pressure, presumably by shifting the equilibrium between surface and bulk molecules.

At this point we may ask, why should water-soluble molecules leave an aqueous phase and accumulate or "adsorb" at an air–solution interface? Since any process will occur spontaneously if it results in a net loss in free energy, such must be the case for the process of adsorption.

A number of factors will produce such a favorable change in free energy. First, the presence of the oriented monomolecular film reduces the surface free energy of the air–water interface. Second, the hydrophobic group on the molecule is in a lower state of energy at the interface, where it no longer is as surrounded by water molecules, than when it is in the bulk-solution phase. Increased interaction between film molecules also will contribute to this process.

A further reduction in free energy occurs upon adsorption because of the gain in entropy associated with a change in water structure. Water molecules, in the presence of dissolved alkyl chains are more highly organized or "ice-like" than they are as a pure bulk phase; hence, the entropy of such structured water is lower than that of bulk water.

The process of adsorption requires that the "ice-like" structure "melt" as the chains go to the interface and, thus, an increase in the entropy of water occurs. The adsorption of molecules dissolved in oil can occur but it is not influenced by water structure changes and, hence, only the first factors mentioned are important here.

It is very rare that significant adsorption can occur at the hydrocarbon–air interface since little loss in free energy can occur by bringing hydrocarbon chains with polar groups attached to this interface; however, at oil–water interfaces the polar portions of the molecule can interact with water at the interface, leading to significant adsorption.

Thus, whereas water-soluble fatty acid salts are adsorbed from water to air–water and oil–water interfaces, their undissociated counterparts, the free fatty acids, which are water insoluble, form insoluble films at the air–water interface, are not adsorbed from oil solution to an oil–air interface, but show significant adsorption at the oil–water interface when dissolved in oil.

From this discussion it is possible also to conclude that adsorption from aqueous solution requires a lower solute concentration to obtain the same level of adsorption if the hydrophobic chain length is increased or if the polar portion of the molecule is less hydrophilic. On the other hand, adsorption from nonpolar solvents is favored when the solute is quite polar.

Since soluble or adsorbed films cannot be compressed, there is no simple direct way to estimate the number of molecules/unit area coming to the surface under a given set of conditions. For relatively simple systems it is possible to estimate this value by application of the Gibbs equation, which relates surface concentration to the surface-tension change produced at different solute activities. The derivation of this equation is beyond the scope of this discussion, but it arises from a classical thermodynamic treatment of the change in free energy when molecules concentrate at the boundary between two phases. The equation may be expressed as

$$\Gamma = -\frac{a}{RT}\frac{d\gamma}{da} \tag{29}$$

where Γ is the moles of solute adsorbed/unit area, R is the gas constant, T is the absolute temperature and $d\gamma$ is the change in surface tension with a change in solute activity, da, at activity a. For dilute solutions of nonelectrolytes, or for electrolytes when the Debye–Hückel equation for activity coefficient is applicable, the value of a may be replaced by solute concentration, c. Since the term dc/c is equal to $d \ln c$, the Gibbs equation is often written as

$$\Gamma = -\frac{1}{RT}\frac{d\gamma}{d \ln c} \tag{30}$$

In this way the slope of a plot of γ vs $\ln c$ multiplied by $1/RT$ should give Γ at a particular value of c. Figure 19-13 depicts typical plots for a series of water-soluble surface-active agents differing only in the alkyl chain length. Note the

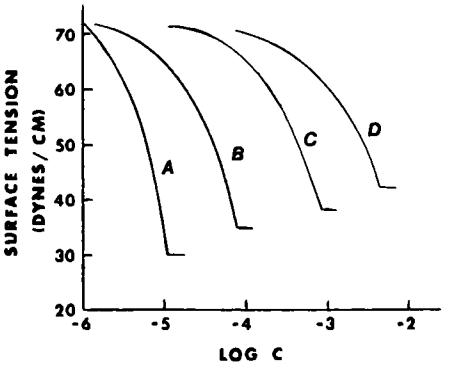


Fig 19-13. The effect of increasing chain length on the surface activity of a surfactant at the air–aqueous solution interface (each figure depicted to differ by two methylene groups with *A*, the longest chain, and *D*, the shortest).

greater reduction of surface tension that occurs at lower concentrations for longer chain-length compounds. In addition, note the greater slopes with increasing concentration, indicating more adsorption (Eq 30), and the abrupt leveling of surface tension at higher concentrations. This latter behavior reflects the self-association of surface-active agent to form micelles which exhibit no further tendency to reduce surface tension. The topic of micelles will be discussed later on page 268.

If one plots the values of surface concentration, Γ, vs concentration, c, for substances adsorbing to the vapor-liquid and liquid-liquid interfaces, using data such as those given in Fig 19-13, one generally obtains an adsorption isotherm shaped like those in Fig 19-9 for vapor adsorption. Indeed, it can be shown that the Langmuir equation (Eq 25) can be fitted to such data when written in the form

$$\Gamma = \frac{\Gamma_{max} k'c}{1 + k'c} \tag{31}$$

where Γ_{max} is the maximum surface concentration attained with increasing concentration and k' is related to k in Eq 25. Combining Eqs 29 and 31 leads to a widely used relationship between surface tension change Π (see Eq 28) and solute concentration, c, known as the Syszkowski equation:

$$\Pi = \Gamma_{max} RT \ln (1 + k'c) \tag{32}$$

Mixed Films

It would seem reasonable to expect that the properties of a surface film could be varied greatly if a mixture of surface-active agents were in the film. As an example, consider that a mixture of short- and long-chain fatty acids would be expected to show a degree of "condensation" varying from the "gaseous" state, when the short-chain substance is used in high amount, to a highly condensed state when the longer chain substance predominates. Thus, each component in such a case would operate independently by bringing a proportional amount of film behavior to the system.

More often, the ingredients of a surface film do not behave independently, but, rather, interact to produce a new surface film. An obvious example would be the combination of organic amines and acids which are oppositely charged and would be expected to interact strongly.

In addition to such polar-group interactions, chain–chain interaction will strongly favor mixed condensed films. An important example of such a case occurs when a long-chain alcohol is introduced along with an ionized long-chain substance. Together the molecules form a highly condensed film despite the presence of a high number of like charges. Presumably this occurs as seen in Fig 19-14, by arranging the molecules so that ionic groups alternate with alcohol groups; however, if chain–chain interactions are not strong, the ionic species often will be displaced by the more nonpolar unionized species and "desorb" into the bulk solution.

On the other hand, sometimes the more soluble surface-active agent produces surface pressures in excess of the collapse pressure of the insoluble film and displaces it from the surface. This is an important concept because it is the underlying principle behind cell lysis by surface-active agents and some drugs, and behind the important process of detergency.

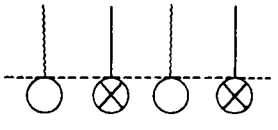


Fig 19-14. A mixed monomolecular film. ⊗: a long-chain ion; O: a long-chain nonionic compound.

Adsorption on Solid Surfaces From Solution

Adsorption to solid surfaces from solution may occur if dissolved molecules and the solid surface have chem groups capable of interacting. Nonspecific adsorption will occur if the solute is surface active and if the surface of the solid is high. This latter case would be the sam occurs at the vapor-liquid and liquid-liquid interfaces. with adsorption to liquid interfaces, adsorption to solid faces from solution generally leads to a monomolecular er, often described by the Langmuir equation or by empirical, yet related, Freundlich equation

$$x/M = kc^n$$

where x is the grams of solute adsorbed by M grams of s in equilibrium with a solute concentration of c. The terr and n are empirical constants. However, as Giles[8] pointed out, the variety of combinations of solutes and ids, and, hence the variety of possible mechanisms of ads tion, can lead to a number of more complex isotherms. particular, adsorption of surfactants and polymers, of g importance in a number of pharmaceutical systems, is not well understood on a fundamental level, and ma some situations even be multilayered.

Adsorption from solution may be measured by separa solid and solution and either estimating the amount of sorbate adhering to the solid or the loss in concentratio adsorbate from solution.

In view of the possibility of solvent adsorption, the la approach really only gives an apparent adsorption. example, if solvent adsorption is great enough, it is poss to end up with an increased concentration of solute a contact with the solid; here, the term negative adsorptic used.

Solvent not only influences adsorption by competing the surface but, as discussed in connection with adsorp at liquid surfaces, the solvent will determine the escap tendency of a solute; eg, the more polar the molecule, the the adsorption that occurs from water. This is seen in 19-15, where adsorption of various fatty acids from w onto charcoal increases with increasing alkyl chain lengt nonpolarity. It is difficult to predict these effects but general, the more chemically unlike the solute and solv and the more alike the solid surface groups and solute,

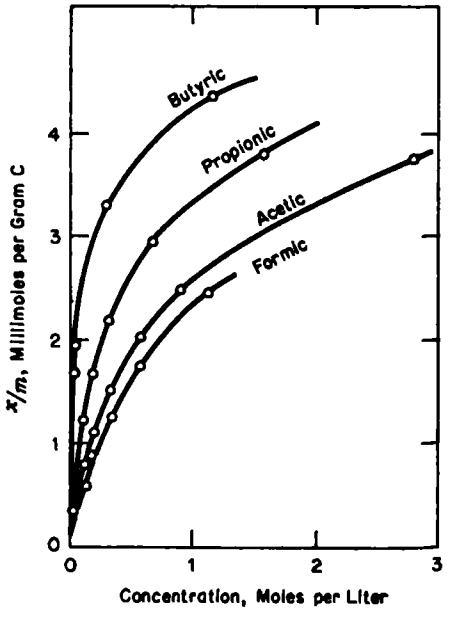


Fig 19-15. The relation between adsorption and molecular weig fatty acids.[9]

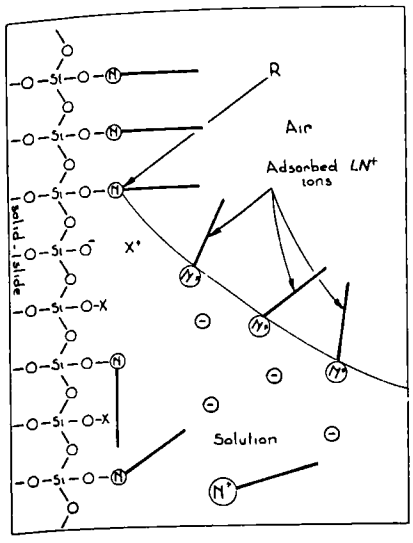


Fig 19-16. The adsorption of a cationic surfactant, LN^+, onto a negatively charged silica or glass surface, exposing a hydrophobic surface as the solid is exposed to air.[10]

greater the extent of adsorption. Another factor which must be kept in mind is that charged solid surfaces, such as polyelectrolytes, will strongly adsorb oppositely charged solutes. This is similar to the strong specific binding seen in gas chemisorption and it is characterized by significant monolayer adsorption at very low concentrations of solute. See Fig 19-16 for an example of such adsorption.

Surface-Active Agents

Throughout the discussion so far, examples of surface-active agents (surfactants) have been restricted primarily to fatty acids and their salts. It has been shown that both a hydrophobic portion (alkyl chain) and a hydrophilic portion (carboxyl and carboxylate groups) are required for their surface activity, the relative degree of polarity determining the tendency to accumulate at interfaces. It now becomes important to look at some of the specific types of surfactants available and to see what structural features are required for different pharmaceutical applications.

The classification of surfactants is quite arbitrary, but one based on chemical structure appears best as a means of introducing the topic. It is generally convenient to categorize surfactants according to their polar portions since the nonpolar portion is usually made up of alkyl or aryl groups. The major polar groups found in most surfactants may be divided as follows: anionic, cationic, amphoteric and nonionic. As we shall see, the last group is the largest and most widely used for pharmaceutical systems, so that it will be emphasized in the discussion that follows.

Types

Anionic Agents—The most commonly used anionic surfactants are those containing carboxylate, sulfonate, and sulfate ions. Those containing carboxylate ions are known as soaps and are generally prepared by the saponification of natural fatty acid glycerides in alkaline solution. The most common cations associated with soaps are sodium, potassium, ammonium, and triethanolamine, while the chain length of the fatty acids ranges from 12 to 18.

The degree of water solubility is greatly influenced by the length of the alkyl chain and the presence of double bonds. For example, sodium stearate is quite insoluble in water at room temperature, whereas sodium oleate under the same conditions is quite water soluble.

Table VII—Effect of Aerosol OT Concentration on the Surface Tension of Water and the Contact Angle of Water with Magnesium Stearate

Concentration, $m \times 10^6$	γ_{sv}	θ
1.0	60.1	120°
3.0	49.8	113°
5.0	45.1	104°
8.0	40.6	89°
10.0	38.6	80°
12.0	37.9	71°
15.0	35.0	63°
20.0	32.4	54°
25.0	29.5	50°

Multivalent ions, such as calcium and magnesium, produce marked water insolubility, even at lower alkyl chain lengths; thus, soaps are not useful in hard water which is high in content of these ions. Soaps, being salts of weak acids, are subject also to hydrolysis and the formation of free acid plus hydroxide ion, particularly when in more concentrated solution.

To offset some of the disadvantages of soaps, a number of long-alkyl-chain sulfonates, as well as alkyl aryl sulfonates such as sodium dodecylbenzene sulfonate, may be used; the sulfonate ion is less subject to hydrolysis and precipitation in the presence of multivalent ions. A popular group of sulfonates, widely used in pharmaceutical systems, are the dialkyl sodium sulfosuccinates, particularly sodium bis-(2-ethylhexyl)sulfosuccinate, best known as Aerosol OT or docusate sodium. This compound is unique in that it is both oil and water soluble and hence forms micelles in both phases. It reduces surface and interfacial tension to low values and acts as an excellent wetting agent in many types of solid dosage forms (see Table VII).

A number of alkyl sulfates are available as surfactants, but by far the most popular member of this group is sodium lauryl sulfate, which is widely used as an emulsifier and solubilizer in pharmaceutical systems. Unlike the sulfonates, sulfates are susceptible to hydrolysis which leads to the formation of the long-chain alcohol, so that pH control is most important for sulfate solutions.

Cationic Agents—A number of long-chain cations, such as amine salts and quaternary ammonium salts, are often used as surface-active agents when dissolved in water; however, their use in pharmaceutical preparations is limited to that of antimicrobial preservation rather than as surfactants. This arises because the cations adsorb so readily at cell membrane structures in a nonspecific manner, leading to cell lysis (eg, hemolysis), as do anionics to a lesser extent. It is in this way that they act to destroy bacteria and fungi.

Since anionic and nonionic agents are not as effective as preservatives, one must conclude that the positive charge of these compounds is important; however, the extent of surface activity has been shown to determine the amount of material needed for a given amount of preservation. Quaternary ammonium salts are preferable to free amine salts since they are not subject to effect by pH in any way; however, the presence of organic anions such as dyes and natural polyelectrolytes is an important source of incompatibility and such a combination should be avoided.

Amphoteric Agents—The major group of molecules falling into this category are those containing carboxylate or phosphate groups as the anion and amino or quaternary ammonium groups as the cation. The former group is represented by various polypeptides, proteins, and the alkyl betaines, while the latter group consist of natural phospholipids such as the lecithins and cephalins. In general, long-chain amphoterics which exist in solution in zwitterionic form are

more surface-active than ionic surfactants having the same hydrophobic group since in effect the oppositely charged ions are neutralized. However, when compared to nonionics, they appear somewhere between ionic and nonionic.

Nonionic Agents—The major class of compounds used in pharmaceutical systems are the nonionic surfactants since their advantages with respect to compatibility, stability, and potential toxicity are quite significant. It is convenient to divide these compounds into those that are relatively water insoluble and those that are quite water soluble.

The major type of compounds making up this first group are the long-chain fatty acids and their water-insoluble derivatives. These include (1) fatty alcohols such as lauryl, cetyl (16 carbons) and stearyl alcohols; (2) glyceryl esters such as the naturally occurring mono-, di- and triglycerides; and (3) fatty acid esters of fatty alcohols and other alcohols such as propylene glycol, polyethylene glycol, sorbitan, sucrose and cholesterol. Included also in this general class of nonionic water-insoluble compounds are the free steroidal alcohols such as cholesterol.

To increase the water solubility of these compounds and to form the second group of nonionic agents, polyoxyethylene groups are added through an ether linkage with one of their alcohol groups. The list of derivatives available is much too long to cover completely, but a few general categories will be given.

The most widely used compounds are the polyoxyethylene sorbitan fatty acid esters which are found in both internal and external pharmaceutical formulations. Closely related compounds include polyoxyethylene glyceryl, and steroidal esters, as well as the comparable polyoxypropylene esters. It is also possible to have a direct ether linkage with the hydrophobic group as with a polyoxyethylene–stearyl ether or a polyoxyethylene–alkyl phenol. These ethers offer advantages since, unlike the esters, they are quite resistant to acidic or alkaline hydrolysis.

Besides the classification of surfactants according to their polar portion, it is useful to have a method that categorizes them in a manner that reflects their interfacial activity and their ability to function as wetting agents, emulsifiers, solubilizers, etc. Since variation in the relative polarity or nonpolarity of a surfactant significantly influences its interfacial behavior, some measure of polarity or nonpolarity should be useful as a means of classification.

One such approach assigns a hydrophile–lipophile balance number (HLB) for each surfactant and, although developed by a commercial supplier of one group of surfactants, the method has received wide-spread application. The HLB value, as originally conceived for nonionic surfactants, is merely the percentage weight of the hydrophilic group divided by five in order to reduce the range of values. On a molar basis, therefore, a 100% hydrophilic molecule (polyethylene glycol) would have a value of 20.

Thus, an increase in polyoxyethylene chain length increases polarity and, hence, the HLB value; at constant polar chain length, an increase in alkyl chain length or number of fatty acid groups decreases polarity and the HLB value. One immediate advantage of this system is that to a first approximation one can compare any chemical type of surfactant to another type when both polar and nonpolar groups are different.

HLB values for nonionics are calculable on the basis of the proportion of polyoxyethylene chain present; however, in order to determine values for other types of surfactants it is necessary to compare physical chemical properties reflecting polarity with those surfactants having known HLB values.

Relationships between HLB and phenomena such as water solubility, interfacial tension, and dielectric constant have been used in this regard. Those surfactants exhibiting values greater than 20 (eg, sodium lauryl sulfate) demonstrate hydrophilic behavior in excess of the polyoxyethylene groups alone. Table XIX, page 304, presents HLB values for a variety of surface-active agents.

Surfactant Properties in Solution and Micelle Formation

As seen in Fig 19-13, increasing the concentration of surface-active agents in aqueous solution causes a decrease in the surface tension of the solution until a certain concentration where it then becomes essentially constant with increasing concentration. That this change is associated with changes also taking place in the bulk solution rather than just at the surface can be seen in Fig 19-17, which shows the same abrupt change in bulk solution properties such as solubility, equivalent conductance and osmotic pressure as with surface properties. The most reasonable explanation for these effects is that the solute molecules self-associate to form soluble aggregates which exhibit markedly different properties from the monomers in solution. Such aggregates (Fig 19-18*A*) appear to exhibit no tendency to adsorb to the surface since the surface and interfacial tension above this solute concentration do not change to any significant extent. Such aggregates, known as micelles, form over such a very narrow range of concentrations that one can speak of a critical micellization concentration (cmc). These micelles form for essentially the same reasons that cause molecules to be adsorbed; the lack of affinity of the hydrophobic chains for water molecules and the tendency for strong hydrophobic chain-chain interactions when the chains are oriented closely together in the micelle, coupled with the gain in entropy due to the loss of the ice-like structure of water when the chains are separated from water, lead to a favorable free energy change for micellization. The longer the hydrophobic chain or the less the polarity of the polar group, the greater the tendency for monomers to "escape" from the water to form micelles and, hence the lower the cmc (see Fig 19-13).

In dilute solution (still above the cmc) the micelles can be considered to be approximately spherical in shape (Fig 19-18*A* and *B*), while at higher concentrations they become

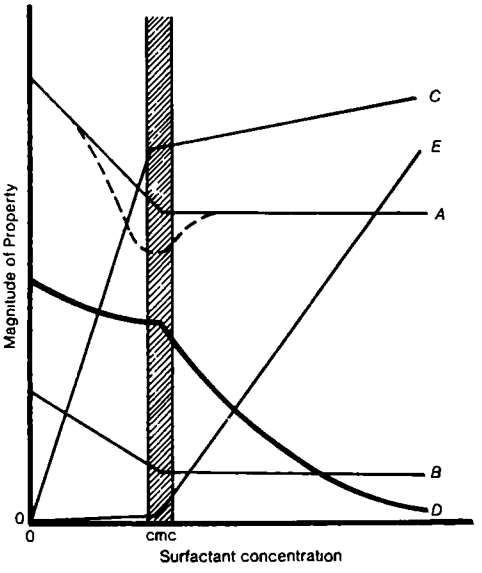


Fig 19-17. Effect of surfactant concentration and micelle formation on various properties of the aqueous solution of an ionic surfactant. *A:* Surface tension; *B:* interfacial tension; *C:* osmotic pressure; *D:* equivalent conductivity; *E:* solubility of compound with very low solubility in pure water [11]

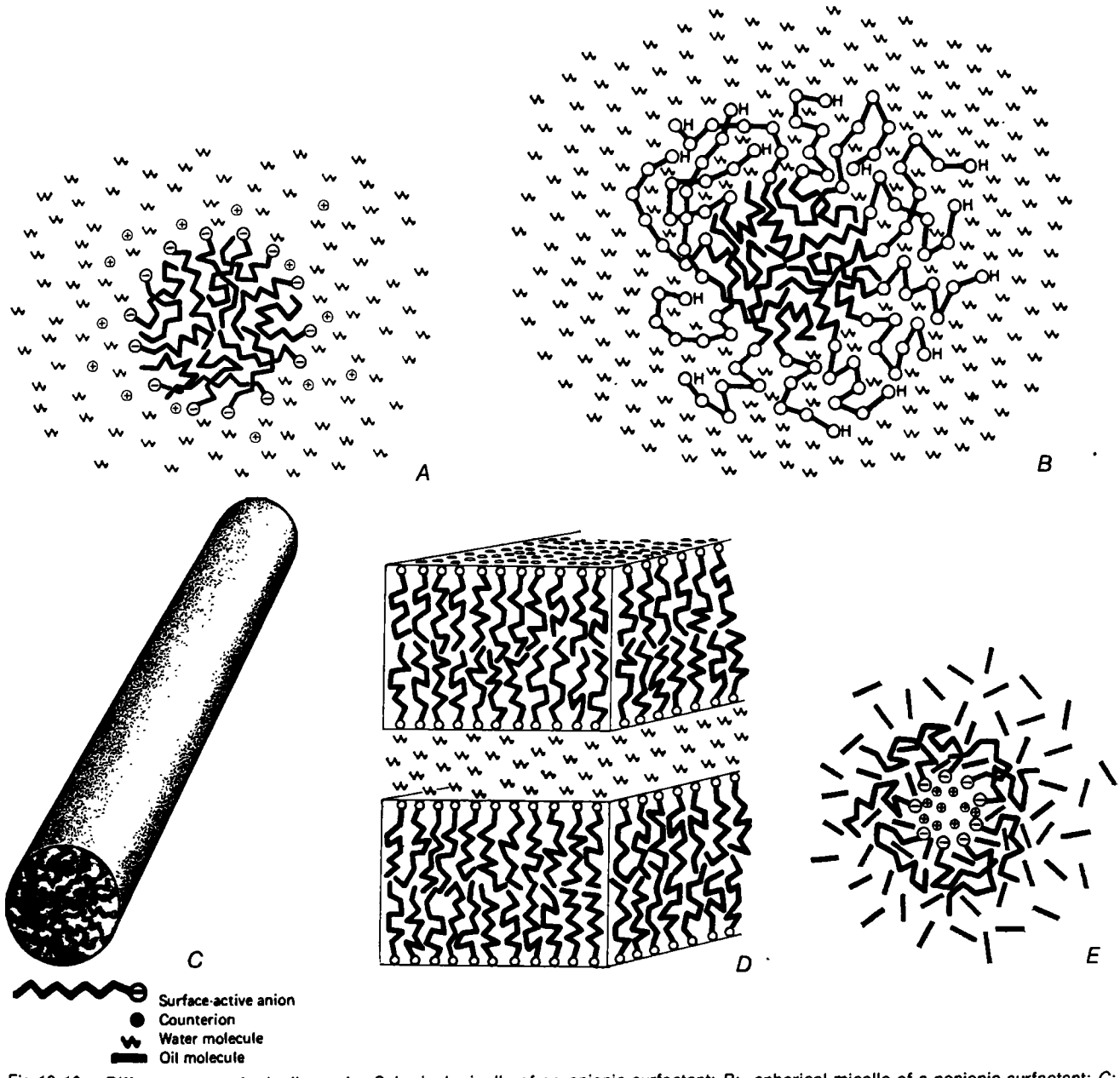


Fig 19-18. Different types of micelles. *A:* Spherical micelle of an anionic surfactant; *B:* spherical micelle of a nonionic surfactant; *C:* cylindrical micelle of an ionic surfactant; *D:* lamellar micelle of an ionic surfactant; *E:* reverse micelle of an anionic surfactant in oil.[11]

more asymmetric and eventually assume cylindrical (Fig 19-18*C*) or lamellar (Fig 19-18*D*) structures. It is important to recognize that equilibrium, and hence reversibility, exists between the monomers and the various types of micelles. The sizes of such micelles depend on the number of monomers per micelle and the size and molecular shape of the individual monomers. In Table VIII are given the cmc and number of monomers per micelle for different types of surfactants. Note for the nonionic surfactants that the longer the polyoxyethylene chain, and hence the more polar and bulkier the molecule, the higher the cmc, ie the less the tendency for micelle formation. It is also possible for oil-soluble surfactants to show a tendency to self-associate into "reverse micelles in nonpolar solvents, as depicted in Fig 19-18*E*, with their polar groups all oriented away from the solvent. In general these micelles tended to be smaller and to aggregate over a wider range of concentrations than seen in water, and therefore, to exhibit no well-defined cmc.

Micellar Solubilization

As seen in Fig 19-18, the interior of surfactant micelles formed in aqueous media consists of hydrocarbon "tails" in liquid-like disorder. The micelles, therefore, resemble miniscule pools of liquid hydrocarbon surrounded by shells of polar "head groups." Compounds which are poorly soluble in water but soluble in hydrocarbon solvents, can be dissolved inside these micelles, ie, they are brought homogeneously into an overall aqueous medium.

Being hydrophobic and oleophilic, the solubilized molecules are located primarily in the hydrocarbon core of the micelles (see Fig 19-19*A*). Even water-insoluble drugs usually contain polar functional groups such as hydroxyl, carbonyl, ether, amino, amide, and cyano. Upon solubilization, these hydrophilic groups locate on the periphery of the micelle among the polar headgroups of the surfactant in order to become hydrated (see Fig 19-19*B*). For instance,

Table VIII—Critical Micelle Concentrations and Micellar Aggregation Numbers of Various Surfactants in Water at Room Temperature

Structure	Name	CMC, m*M*/L	Surfactant molecules/micelle
n-$C_{11}H_{23}COOK$	Potassium laurate	24	50
n-$C_8H_{17}SO_3Na$	Sodium octant sulfonate	150	28
n-$C_{10}H_{21}SO_3Na$	Sodium decane sulfonate	40	40
n-$C_{12}H_{25}SO_3Na$	Sodium dodecane sulfonate	9	54
n-$C_{12}H_{25}OSO_3Na$	Sodium lauryl sulfate	8	62
n-$C_{12}H_{25}OSO_3Na$	Sodium lauryl sulfate[a]	1	96
	Sodium di-2-ethylhexyl sulfosuccinate	5	48
n-$C_{10}H_{21}N(CH_3)_3Br$	Decyltrimethylammonium bromide	63	36
n-$C_{12}H_{25}N(CH_3)_3Br$	Dodecyltrimethylammonium bromide	14	50
n-$C_{14}H_{29}N(CH_3)_3Br$	Tetradecyltrimethylammonium bromide	3	75
n-$C_{14}H_{29}N(CH_3)_3Cl$	Tetradecyltrimethylammonium chloride	3	64
n-$C_{12}H_{25}NH_3Cl$	Dodecylammonium chloride	13	55
n-$C_{12}H_{25}O(CH_2CH_2O)_8H$	Octaoxyethylene glycol monododecyl ether	0.13	132
n-$C_{12}H_{25}O(CH_2CH_2O)_8H$[b]		0.10	301
n-$C_{12}H_{25}(CH_2CH_2O)_{12}H$	Dodecaoxyethylene glycol monododecyl ether	0.14	78
n-$C_{12}H_{25}O(CH_2CH_2O)_{12}H$[b]		0.091	116
t-C_8H_{17}-C_6H_4-$O(CH_2CH_2O)_{9.7}H$	Decaoxyethylene glycol mono-p,t-octylphenyl ether (octoxynol 9)	0.27	100

[a] Interpolated for physiologic saline, 0 154 *M* NaCl.
[b] At 55° instead of 20°.

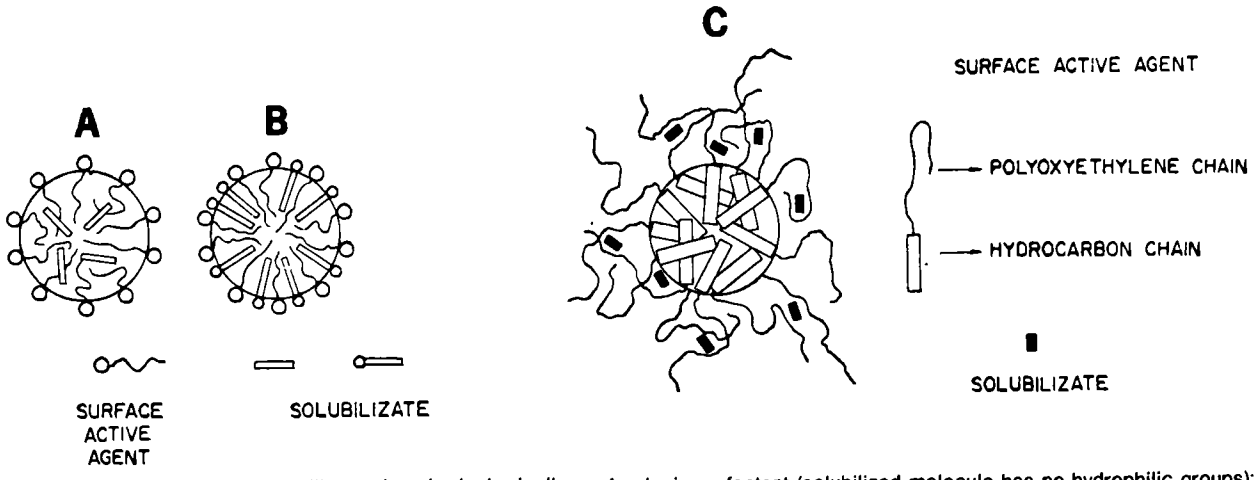


Fig 19-19. The locations of solubilizates in spherical micelles. A: Ionic surfactant (solubilized molecule has no hydrophilic groups); B: ionic surfactant (solubilized molecule has a hydrophilic group); C: nonionic surfactant (polar solubilizate) [12]

when cholesterol or dodecanol is solubilized by sodium lauryl sulfate, their hydroxyl groups penetrate between sulfate ions and are even bound to them by hydrogen bonds, while their hydrocarbon portions are immersed among the dodecyl tails of the surfactant which make up the core of the micelle.

Micelles of polyoxyethylated nonionic surfactants consist of an outer shell of hydrated polyethylene glycol moieties and a core of hydrocarbon moieties. Compounds like phenol, cresol, benzoic acid, salicylic acid, and esters of *p*-hydroxy and *p*-aminobenzoic acids have some solubility in water and in oils but considerable solubility in liquids of intermediate polarity like ethanol, propylene glycol or aqueous solutions of polyethylene glycols. When solubilized by nonionic micelles, they are located in the hydrated outer polyethylene glycol shell as shown in Fig 19-19C. Since these compounds have hydroxyl or amino groups, they frequently form complexes with the ether oxygens of the surfactant by hydrogen bonding.

Solubilization is generally nonspecific: any drug which is appreciably soluble in oils can be solubilized. Each has a solubilization limit, comparable to a limit of solubility, which depends on temperature and on the nature and concentration of the surfactant. Hartley distinguishes two categories of solubilizates. The first consists of comparatively large, asymmetrical and rigid molecules forming crystalline solids, such as steroids and dyes. These do not blend in with the normal paraffin tails which make up the micellar core; because of dissimilarity in structure, they remain distinct as solute molecules. They are sparingly solubilized by surfactant solutions, a few molecules/micelle at saturation (see Table IX). The number of carbon atoms in the micellar hydrocarbon core required to solubilize a molecule of steroid or dye at saturation is of the same order of magnitude as the number of carbon atoms of bulk liquid dodecane or hexadecane per molecule of steroid or dye in their saturated solutions in these liquids.

Since solubilization depends on the presence of micelles, it does not take place below the cmc. It can, therefore, be used to determine the cmc, particularly when the solubilizate is a dye or another compound easy to assay. Plotting the maximum amount of a water-insoluble dye solubilized by aqueous surfactant, or the absorbance of its saturated solutions, versus the surfactant concentration produces a straight line which intersects the surfactant concentration axis at the cmc. Above the cmc, the amount of solubilized dye is directly proportional to the number of micelles and, therefore,

Table IX—Micellar Solubilization Capacities of Different Surfactants for Estrone[13]

Surfactant	Concentration range, molarity	Temp, °C	Moles surfactant/ mole solubilized estrone
Sodium laurate	0.025–0.023	40	91
Sodium oleate	0.002–0.35	40	53
Sodium lauryl sulfate	0.004–0.15	40	71
Sodium cholate	0.09–0.23	20	238
Sodium deoxycholate	0.007–0.36	20	476
Diamyl sodium sulfosuccinate	0.08–0.4	40	833
Dioctyl sodium sulfosuccinate	0.002–0.05	40	196
Tetradecyltrimethylammonium bromide	0.005–0.08	20	45
Hexadecylpyridinium chloride	0.001–0.1	20	32
Polysorbate 20	0.002–0.15	20	161
Polysorbate 60	0.0008–0.11	20	83

proportional to the overall surfactant concentration. Below the cmc, no solubilization takes place. This is represented by Curve E of Fig 19-17.

The second category of compounds to be solubilized are often liquid at room temperature and consist of relatively small, symmetrical, and/or flexible molecules such as many constituents of essential oils. These molecules mix and blend in freely with the hydrocarbon portions of the surfactants in the core of the micelles, so as to become indistinguishable from them. Such compounds are extensively solubilized and in the process usually swell the micelles: they augment the volume of the hydrocarbon core and increase the number of surfactant molecules per micelle. Their solubilization frequently lowers the cmc.

Microemulsions[14-16]

Microemulsions are liquid dispersions of water and oil that are made homogeneous, transparent, and stable by the addition of relatively large amounts of a surfactant and a cosurfactant. *Oil* is defined as a liquid of low polarity and low miscibility with water, eg, toluene, cyclohexane, mineral or vegetable oils.

Microemulsions are intermediate in properties between micelles containing solubilized oils and emulsions. While emulsions are lyophobic and unstable, microemulsions are on the borderline between lyophobic and lyophilic colloids. True microemulsions are thermodynamically stable.[17] Therefore, they are formed spontaneously when oil, water, surfactants, and cosurfactants are mixed together. The unstable emulsions require input of considerable mechanical energy for their preparation, which may be supplied by colloid mills, homogenizers or ultrasonic generators.

Both emulsions and microemulsions may contain high volume fractions of the internal phase. For instance, some O/W systems contain 75% (v/v) of oil dispersed in 25% water, although lower internal phase volume fractions are more common.

At low surfactant concentrations, viz, low multiples of the cmc, micelles are spheres (Fig 19-18*A*, *B* and *E*) or ellipsoids. When an oil is solubilized by micelles in water, it blends into the micellar core formed by the hydrocarbon tails of the surfactant molecules (Fig 19-19) and swells the micelles.

Spherical or ellipsoidal micelles are nearly monodisperse, and their mean diameters are in the range of 25 to 60 Å. Microemulsion droplets also have a narrow droplet size distribution with a mean diameter range of approximately 60 to 1000 Å. Since the droplet diameters are less than ¼ of the wavelength of light (4200 Å for violet and 6600 Å for red light), microemulsions scatter little light and are, therefore, transparent or at least translucent.

Emulsions have very broad droplet size distributions. Only the smallest droplets, with diameters of about 1000 to 2000 Å, are below the resolving power of the light microscope. The upper size limit is 25 or 50 μm (250,000 or 500,000 Å). Because emulsion droplets are comparable in size, or larger than the wavelength of visible light, they scatter it more or less strongly depending on the difference in refractive index between oil and water. Thus, most emulsions are opaque.

The three disperse systems—micellar solutions, microemulsions, and emulsions—can be of the O/W (oil-in-water) or W/O type. Aqueous micellar surfactant solutions can solubilize oils and lipid-soluble drugs in the core formed by their hydrocarbon chains. Likewise, oil-soluble surfactants like sorbitan monooleate and docusate sodium form "reverse micelles" in oils (Fig 19-18*E*) capable of solubilizing water in the polar center. The solubilized oil in the former micelles and the solubilized water in the latter may in turn enhance the micellar solubilization of oil-soluble and water-soluble drugs, respectively.

Oil-soluble drugs have been incorporated into O/W emulsions by dissolving them in the oil phase before emulsification.[18] By the same token, it may be possible to dissolve oil-soluble drugs in a vegetable oil and make an oral or parenteral O/W microemulsion. The advantage of such microemulsion systems over conventional emulsions is their smaller droplet size and superior shelf stability. Aqueous micellar solutions[19] and O/W microemulsions[20] have both been used as aqueous reaction media for oil-soluble compounds.

Emulsions and micellar solutions of oils solubilized in aqueous surfactant solutions consist of three components, oil, water and surfactant. Microemulsions generally require a fourth component, called *cosurfactant*. Commonly used cosurfactants are linear alcohols of medium chain length, which are sparingly miscible with water. Since the cosurfactants as well as the surfactants are surface-active, they promote the generation of extensive interfaces through the spontaneous dispersion of oil in water, or vice-versa, resulting in the formation of microemulsions. The large interfacial area between oil and water permits the extensive formation of a mixed interfacial film consisting of surfactant and cosurfactant. This film is called the "interphase" because it is thicker than the surfactant monolayers formed at oil-water interfaces in emulsions. The interfacial tension at the oil-water interface in microemulsions approaches zero, which also contributes to their spontaneous formation. According to another viewpoint, microemulsions are regarded as micelles extensively swollen by large amounts of solubilized oil.

Typical formulations for an O/W and a W/O microemulsion are shown in Table X. The ratio, g surfactant/g solubilized or emulsified oil or water is in the range of 2 to 20 for micellar solutions and 0.01 to 0.1 for emulsions. Microemulsions have intermediate values: The ratios for the formulations in Table X are near unity. In industrial formulations,

Table X—Microemulsion Formulations

Compound	Function	Content in microemulsions, % O/W	Content in microemulsions, % W/O
Sodium lauryl sulfate	Surfactant	13	10
1-Pentanol	Cosurfactant	8	25
Xylene	Oil	8	50
Water		71	15

the ratios are closer to 0.1 to reduce costs. Microemulsions are used in such diverse applications as floor polish and agricultural pesticide formulations and in tertiary petroleum recovery. The use of O/W microemulsions as aqueous vehicles for oil-soluble drugs to be administered by the percutaneous, oral or parenteral route is being investigated.

Colloidal Dispersions

Historical Background of Colloids

The term *colloid*, derived from the Greek word for glue, was applied *ca* 1850 by the British chemist Thomas Graham to polypeptides such as albumin and gelatin, to vegetable gums such as acacia, starch and dextrin, and to inorganic compounds such as gelatinous metal hydroxides and Prussian blue (ferric ferrocyanide). These compounds did not crystallize, and diffused very slowly when dissolved or dispersed in water. They could be separated from ordinary solutes such as salts and sugar, called "crystalloids," as the latter diffused through the fine pores of dialysis membranes made from animal gut which retained the "colloids." "Crystalloids" crystallized readily from solution.[21,22]

Von Weimarn was the first to identify colloidality as a state of subdivision of matter rather than as a category of substances. Many of Graham's "colloids," especially proteins, have been crystallized. Moreover, von Weimarn was able to prepare all "crystalloids" investigated in the colloidal state. Colloidal dispersions by the condensation method resulted from high relative supersaturation, which produced a large number of small nuclei.[21-23,28] For instance, clear, transparent solidified jellies were prepared by cooling aqueous solutions of $CaCl_2$, $Ba(SCN)_2$ and $Al_2(SO_4)_3$, and aqueous-alcoholic solutions of NaCl, KCl, NH_4Cl, KSCN, NaBr and NH_4NO_3 which were nearly saturated at room temperature.[28]

Colloid chemistry became a science in its own right around 1906, when Wolfgang Ostwald wrote the booklet "The World of the Neglected Dimensions." In it, he focused on colloidal systems as a state of matter that has disperse phases intermediate in size between small molecules or ions in solution and large, visible particles in suspension. Ostwald became the first editor of the journal *Kolloid-Zeitschrift* in 1907. The studies of colloidal systems and surface or interfacial phenomena are intimately related. The properties of colloidal dispersions are largely governed by the nature of the surface of their particles. The division of the American Chemical Society specializing in colloidal systems and interfaces is called the "Division of Colloid and Surface Chemistry," while the pertinent session of the Gordon Research Conferences is called "Chemistry at Interfaces."

Colloid and surface chemistry deals with an unusually wide variety of industrial and biological systems. A few examples are catalysts, lubricants, adhesives, latexes for paints, rubbers and plastics, soaps and detergents, clays, packaging films, cigarette smoke, liquid crystals, cell membranes, mucous secretions and aqueous humors.

Definitions and Classifications

Colloidal Systems and Interfaces

Colloidal dispersions consist of at least two discrete phases, namely, one or more disperse, dispersed or internal phases and a continuous or external phase called the *dispersion medium* or *vehicle*. What distinguishes colloidal dispersions from solutions and coarse dispersions is the particle size of the disperse phase. Systems in the colloidal state contain one or more substances that have at least one dimension in the range of 10 to 100 Å (1 Angstrom unit = 10^{-8} cm = 10^{-10} m) or 1–10 nm (1 nanometer = 10^{-9} m) at the lower end, and a few micrometers (μm) at the upper end (1 μm = 10^4 Å = 10^{-6} m). Thus blood, cell membranes, the thinner nerve fibers, milk, rubber latex, fog and beer foam are colloidal systems. Some types of materials, such as many emulsions, and oral suspensions of most organic drugs, are coarser than true colloidal systems but exhibit similar behavior. Even though serum albumin, acacia and povidone form true or molecular solutions in water, the size of the individual solute molecules places such solutions in the colloidal range (particle size > 10 Å).[21-27]

The following features distinguish colloidal dispersions from coarse suspensions. Disperse particles in the colloidal range are usually too fine to be visible in a light microscope, because at least one dimension measures 1 μm or less. They are often visible in the ultramicroscope and always in the electron microscope. Coarse suspended particles are frequently visible to the naked eye and always in the light microscope. Colloidal particles, as opposed to coarse particles, pass through ordinary filter paper but are retained by dialysis or ultrafiltration membranes. Because of their small size, colloidal dispersions undergo little or no sedimentation or creaming: Brownian motion maintains the disperse particles in suspension (see below).

Except for high polymers, most soluble substances can be prepared either as low-molecular-weight solutions, or as colloidal dispersions or coarse suspensions depending on the choice of the dispersion medium and the dispersion technique.[26,28]

Because of the small size of colloidal particles, appreciable fractions of their atoms, ions or molecules are located in the boundary layer between a particle and air (surface) or between a particle and a liquid or solid (interface). The ions in the surface of a sodium chloride crystal and the water molecules in the surface of a rain drop are subjected to unbalanced forces of attraction, whereas the ions or molecules in the interior of the materials are surrounded by similar ions or molecules on all sides, with balanced force fields. Thus a surface free energy component is added to the total free energy of colloidal particles, which becomes relatively more important as the particles become smaller, ie, as greater fractions of their ions, atoms or molecules are located in their surface or interfacial region. Hence the solubility of very fine solid particles and the vapor pressure of very small liquid droplets are larger than the corresponding values of coarse particles and large drops of the same materials, respectively.

Specific Surface Area—Decreasing particle size increases the surface-to-volume ratio, which is expressed as the specific surface area A_{sp}, namely, the area A (cm^2) per unit volume V (1 cm^3) or per unit mass M (1 gram). For a sphere, $A = 4\pi r^2$ and $V = 4/3\ \pi r^3$. If the density, d, of the material is expressed in g/cm^3, the specific surface area is

$$A_{sp} = \frac{A}{V} = \frac{4\pi r^2}{4/3\pi r^3} = \frac{3}{r}\ cm^2/cm^3 = \frac{3}{r}\ cm^{-1}$$

or

$$A_{sp} = \frac{A}{M} = \frac{A}{Vd} = \frac{4\pi r^2}{4/3\pi r^3 d} = \frac{3}{rd}\ cm^2/g$$

Table XI—Effect of Comminution on Specific Surface Area of a Volume of $4\pi/3$ cm^3, Divided into Uniform Spheres of Radius R

Number of spheres	R	A_{sp} cm^2/cm^3
1	1 cm	3
10^3	0.1 cm = 1 mm	3×10
10^6	0.1 mm	3×10^2
10^9	0.01 mm = 10 μm	3×10^3
10^{12}	1 μm	3×10^4
10^{15}	0.1 μm	3×10^5
10^{18}	0.01 μm	3×10^6
10^{21}	10 Å = 1 nm	3×10^7
10^{23}	1 Å	3×10^8

Shaded region corresponds to colloidal particle-size range

Table XI illustrates the effect of comminution on the specific surface area of 4 $\pi/3$ cm^3 of a material consisting initially of one sphere of 1 cm radius. As the material is broken up into an increasingly larger number of smaller and smaller spheres, its specific surface area increases commensurately.

The solid adsorbents activated charcoal and kaolin have specific surface areas of about 6×10^6 cm^2/g and 10^4 cm^2/g, respectively. One gram of activated charcoal, because of its extensive porosity and internal voids, has an area equal to 1/6 acre.

In conclusion, colloidal systems by definition are those polyphasic systems where at least one dimension of the disperse phase measures between 10 or 100 Å and a few micrometers. The term "colloidal" designates a state of matter characterized by submicroscopic dimensions rather than certain substances. Any dispersed substance with the proper dimension or dimensions is in the colloidal state.

Physical States of Disperse and Continuous Phases

A useful classification of colloidal systems (systems in the colloidal particle size range) is based on the state of matter of the disperse phase and the dispersion medium, ie, whether they are solid, liquid or gaseous.[25,27] Table XII summarizes the various combinations and lists examples. A *sol* is the colloidal dispersion of a solid in a liquid or gaseous medium. Prefixes designate the dispersion medium, such as hydrosol, alcosol, aerosol for water, alcohol and air, respectively. Sols are fluid. If the solid particles form bridged structures possessing some mechanical strength, the system is called a gel (hydrogel, alcogel, aerogel).

Interaction Between Disperse Phase and Dispersion Medium

A second useful classification of colloidal dispersions, originated by Ostwald, is based on the affinity or interaction between the disperse phase and the dispersion medium.[2,3,8] It refers mostly to solid-in-liquid dispersions. According to this classification, colloidal dispersions are divided into the two broad categories of lyophilic and lyophobic. Some soluble, low-molecular-weight substances have molecules with both tendencies, forming a third category called association colloids.

Lyophilic Dispersions—Where there is considerable attraction between the disperse phase and the liquid vehicle, ie, extensive solvation, the system is said to be *lyophilic* (solvent-loving). If the dispersion medium is water, the system is said to be *hydrophilic*. Such solids as bentonite, starch, gelatin, acacia and povidone swell, disperse or dissolve spontaneously in water.

Hydrophilic colloidal dispersions can be subdivided further as follows:

True solutions, formed by water-soluble polymers (acacia and povidone).

Gelled solutions, gels or jellies if the polymers are present at high concentrations and/or at temperatures where their water solubility is low. Examples of such hydrogels are relatively concentrated solutions of gelatin and starch, which set to gels on cooling, or of methylcellulose, which gel on heating.

Particulate dispersions, where the solids do not form molecular solutions but remain as discrete though minute particles. Bentonite and microcrystalline cellulose form such hydrosols.

Lipophilic or oleophilic substances have pronounced affinity for oils. Oils are nonpolar liquids consisting mainly of hydrocarbons, with few polar groups and low dielectric constants. Examples are mineral oil, benzene, carbon tetrachloride, vegetable oils (cottonseed or peanut oil) and essential oils (lemon or peppermint oil). Substances which form *oleophilic* colloidal dispersions include polymers like polystyrene and unvulcanized or gum rubber, which dissolve molecularly in benzene, magnesium or aluminum stearate or which dissolve or disperse in cottonseed oil, and activated charcoal, which forms sols or particulate dispersions in all oils.

Because of the high affinity or attraction between the dispersion medium and the disperse phase, lyophilic dispersions form spontaneously when the liquid vehicle is brought into contact with the solid phase. They are thermodynamically stable and reversible, ie, they are easily reconstituted even after the dispersion medium has been removed from the solid phase.[22,24-27]

Table XII—Classification of Colloidal Dispersions According to State of Matter

Disperse Phase	Dispersion Medium (Vehicle)		
	Solid	Liquid	Gas
Solid	Zinc oxide paste (zinc oxide + starch in petrolatum). Toothpaste (dicalcium phosphate or calcium carbonate with sodium carboxymethylcellulose binder). Pigmented plastics (titanium dioxide in polyethylene).	Sols: Bentonite Magma NF. Trisulfapyrimidines Oral Suspension USP. Magnesia and Alumina Oral Suspension USP. Tetracycline Oral Suspension USP.	Solid aerosols: Smoke, dust. Epinephrine Bitartrate Inhalation Aerosol USP. Isoproterenol Sulfate Inhalation Aerosol.
Liquid	Absorption bases (aqueous medium in Hydrophilic Petrolatum USP). Emulsion bases (oil in Hydrophilic Ointment USP). Butter.	Emulsions: Mineral Oil Emulsion USP. Soybean oil in water emulsion for IV feeding. Milk. Mayonnaise.	Liquid aerosols: Mist, fog. Nasal relief sprays (naphazoline hydrochloride solution). Betamethasone Valerate Topical Aerosol USP. Povidone-Iodine Topical Aerosol.
Gas	Solid foams (foamed plastics and rubbers). Pumice.	Foams. Carbonated beverages. Effervescent salts in water.	No colloidal dispersions.

Lyophobic Dispersions—When there is little attraction between the disperse phase and the dispersion medium, the dispersion is said to be *lyophobic* (solvent-hating). *Hydrophobic* dispersions consist of particles that are not hydrated, so that water molecules interact with or attract one another in preference to solvating the particles. They include aqueous dispersions of oleophilic materials such as polystyrene or gum rubber (latex), steroids and other organic lipophilic drugs, paraffin wax, magnesium stearate, and of cottonseed or soybean oil (emulsion). While lipophilic materials are generally hydrophobic, materials like sulfur, silver chloride and gold form hydrophobic dispersions without being lipophilic. Water-in-oil emulsions are lyophobic dispersions in lipophilic vehicles.

Because of the lack of attraction between the disperse and the continuous phase, lyophobic dispersions are intrinsically unstable and irreversible. Their large surface free energy is not lowered by solvation. The dispersion process does not take place spontaneously, and once the dispersion medium has been separated from the disperse phase, the dispersion is not easily reconstituted. The dividing line between hydrophilic and hydrophobic dispersions is not very sharp. For instance, gelatinous hydroxides of polyvalent metals such as $Al(OH)_3$ and $Mg(OH)_2$, and clays such as bentonite and kaolin, possess some characteristics of both.[22,24,27]

Association Colloids—Organic compounds which contain large hydrophobic moieties together with strongly hydrophilic groups in the same molecule are said to be amphiphilic. While the individual molecules are generally too small to bring their solutions into the colloidal size range, they tend to associate in aqueous or oil solutions into micelles (see above). Because micelles are large enough to qualify as colloidal particles, such compounds are called association colloids.

Lyophobic Dispersions

Most of the discussion of lyophobic dispersions deals with hydrophobic dispersions or hydrosols (hydrophobic solids or liquids dispersed in aqueous media) because water is the most widely used vehicle. They comprise aqueous dispersions of insoluble organic and inorganic compounds which usually have low degrees of hydration. Organic compounds which are preponderantly hydrocarbon in nature and possess few hydrophilic or polar groups are insoluble in water and hydrophobic.

Hydrophobic dispersions are intrinsically unstable. The most stable state of such systems contains the disperse phase coalesced into large crystals or drops, so that the specific surface area and surface free energy are reduced to a minimum. Therefore, mechanical, chemical or electrical energy must be supplied to the system to break up the disperse phase into small particles, providing for the increase in surface free energy resulting from the parallel increase in specific surface area. Furthermore, special means must be found to stabilize hydrophobic dispersions, preventing the otherwise spontaneous coalescence or coagulation of the disperse phase after it has been finely dispersed.

Preparation and Purification of Lyophobic Dispersions

Colloidal dispersions are intermediate in size between true solutions and coarse suspensions. They can be prepared by aggregation of small molecules or ions until particles of colloidal dimensions result (condensation methods), or by reducing coarse particles to colloidal dimensions through comminution or peptization (dispersion methods).

Dispersion Methods—The first method, *mechanical disintegration* of solids and liquids into small particles and their dispersion in a fluid vehicle, is frequently carried out by input of mechanical energy via shear or attrition. Equipment such as colloid and ball mills, micronizers and, for emulsions, homogenizers is described in Chapters 83 and 88 and in Ref 29. Dry grinding with inert, water-soluble diluting agents also produces colloidal dispersions. Sulfur hydrosols may be prepared by triturating the powder with urea or lactose followed by shaking with water.

Ultrasonic generators provide exceptionally high concentrations of energy. Successful dispersion of solids by means of ultrasonic waves can only be achieved with comparatively soft materials such as many organic compounds, sulfur, talcum, and graphite. Where fine emulsions are mandatory, such as soybean oil-in-water emulsions used for intravenous feeding, emulsification by ultrasound waves is the method of choice.[29] The formation of aerosols is described in Chapter 92.

It should be reiterated that hydrosols of hydrophobic substances are intrinsically unstable. While mechanical disintegration may break up the disperse phase into colloidal particles, the resultant dispersions tend towards separation of that phase. Recrystallization, coagulation or coalescence causes the disperse particles to become progressively coarser and fewer, ultimately resulting in the separation of a macroscopic phase. To avoid this, stabilizing agents must be added during or shortly after the dispersion process (see below). For instance, lecithin may be used to stabilize soybean oil emulsions.

Peptization is a second method for preparing colloidal dispersions. The term, coined by Graham, is defined as the breaking up of aggregates or secondary particles into smaller aggregates or into primary particles in the colloidal size range. Particles which are not formed of smaller ones are called "primary." Peptization is synonymous with *deflocculation*. It can be brought about by the removal of flocculating agents, usually electrolytes, or by the addition of deflocculating or peptizing agents, usually surfactants, water-soluble polymers or ions which are adsorbed at the particle surface.[24,27]

The mechanisms of the following examples are explained in subsequent sections. When powdered activated charcoal is added to water with stirring, the aggregated grains are broken up only incompletely and the resultant suspension is gray and translucent. The addition of 0.1% or less of sodium lauryl sulfate or octoxynol disintegrates the grains into finely dispersed particles forming a deep black and opaque dispersion. Ferric or aluminum hydroxide freshly precipitated with ammonia can be peptized with small amounts of acids which reduce the pH below the isoelectric points of the hydroxides (see below). Even washing the gelatinous precipitate of $Al(OH)_3$ with water tends to peptize it. In quantitative analysis, the precipitate is therefore washed with dilute solutions of ammonium salts that act as flocculating agents, rather than with water.

Condensation Methods—The preparation of sulfur hydrosols is employed to illustrate condensation or aggregation methods. Sulfur is insoluble in water but somewhat soluble in alcohol. When an alcoholic solution of sulfur is mixed with water, a bluish white colloidal dispersion results. In the absence of added stabilizing agents, the particles tend to agglomerate and precipitate on standing. This technique of dissolving the material in a water-miscible solvent such as alcohol or acetone and producing a hydrosol by precipitation with water is applicable to many organic compounds, and has been used to prepare hydrosols of natural resins like mastic, of stearic acid and of polymers (the so-called pseudolatexes).

For sulfur, another less common physical method is to introduce a current of sulfur vapor into water. Condensation produces colloidal particles. Alternatively, the very fine powder produced by condensing sulfur vapor on cold

solid surfaces (sublimed sulfur or flowers of sulfur) can be dispersed in water by addition of a suitable surfactant to produce a hydrosol.

Chemical methods include the reaction between hydrogen sulfide and sulfur dioxide, eg, by bubbling H_2S into an aqueous SO_2 solution:

$$2\,H_2S + SO_2 \rightarrow 3\,S + 2\,H_2O$$

The same reaction occurs when aqueous solutions containing sodium sulfide and sulfite are acidified with an excess of sulfuric or hydrochloric acid. Another reaction is the decomposition of sodium thiosulfate by sulfuric acid, using either very dilute or very concentrated solutions to obtain colloidally dispersed sulfur:

$$H_2SO_4 + 3\,Na_2S_2O_3 \rightarrow 4\,S + 3\,Na_2SO_4 + H_2O$$

Both reactions also produce pentathionic acid, $H_2S_5O_6$, as a by-product. The preferential adsorption of the pentathionate anion at the surface of the sulfur particles confers a negative electric charge on the particles, stabilizing the sol (see below).[22,26,27] When powdered sulfur is boiled with a slurry of lime, it dissolves with the formation of calcium pentasulfide and thiosulfate. Subsequent acidification produces the colloidal "milk of sulfur," which on washing and drying yields Precipitated Sulfur USP (see Chapter 82).

Sols of ferric, aluminum, chromic, stannic and titanium hydroxides or hydrous oxides are produced by hydrolysis of the corresponding chlorides or nitrates:

$$AlCl_3 + 3\,H_2O \rightleftharpoons Al(OH)_3 + 3\,HCl$$

Hydrolysis is promoted by boiling the solution and/or by adding a base to neutralize the acid formed.

Double decompositions producing insoluble salts can lead to colloidal dispersions. Examples are silver chloride and nickel sulfide:

$$NaCl + AgNO_3 \rightarrow AgCl + NaNO_3$$

$$(NH_4)_2S + NiCl_2 \rightarrow NiS + 2\,NH_4Cl$$

Compare also the preparation of White Lotion, which contains precipitated zinc sulfide and sulfur (Chapter 63). Reducing salts of gold, silver, copper, mercury, platinum, rhodium and palladium with formaldehyde, hydrazine, hydroxylamine, hydroquinone or stannous chloride produces hydrosols of the metals. These are strongly colored, eg, red or blue.[21,22,27]

Radioactive Colloids—Colloidal dispersions containing radioactive isotopes find increasing diagnostic and therapeutic application in nuclear medicine. Radioactive colloids that accumulate in tumors and/or lesions or emboli, indicating their location and size, may be used as diagnostic aids. Radioactive colloids with a particle size of about 300 Å, injected intravenously, locate mainly in the reticuloendothelial systems of liver, spleen and other organs and are used in scintillation imaging. The radiation emitted by the colloids is made visible by stationary or scanning devices which show the location, size and shape of the organ being investigated, as well as any tumors within. Radiocolloids are useful in anticancer radiation therapy because of their low solubility, radiation characteristics, and their ability to accumulate and remain located in certain target organs or tumors.[30]

Colloidal gold Au 198 is made by reducing a solution of gold (^{198}Au) chloride either by treatment with ascorbic acid or by heating with an alkaline glucose solution. Gelatin is added as a protective colloid (see below). The particle size ranges from 50 to 500 Å with a mean of 300 Å. The color of the sol is cherry-red in transmitted light. Violet or blue sols have excessively large particle sizes and should be discarded. Colloidal gold is used as a diagnostic and therapeutic aid (see Chapter 33). The half-life of ^{198}Au is 2.7 days.

Technetium 99m sulfur colloid is prepared by reducing sodium pertechnetate ^{99m}Tc with sodium thiosulfate. The product, a mixture of technetium sulfide and sulfur in the colloidal particle size range, is stabilized with gelatin. It is used chiefly in liver, spleen and bone scanning. Its half-life is 6.0 hour.

Microspheres of gelatin or human serum albumin can be prepared in fairly narrow particle-size ranges from 100–200 Å through 45–55 μm. A variety of β- and γ-emitting radionuclides such as ^{131}I, ^{99m}Tc, ^{113m}In or ^{51}Cr can be incorporated to label the microspheres. Such products have been used to scan heart, brain, urogenital and gastrointestinal tracts, liver, and in pulmonary perfusion and inhalation studies.[30]

Refer to Chapters 32 and 33 for an in-depth discussion of radioisotopes.

Organic compounds that are weak bases, such as alkaloids, are usually much more soluble at lower pH values where they are ionized than at higher pH values where they exist as the free base. Increasing the pH of their aqueous solutions well above their pKa may cause precipitation of the free base. Organic compounds which are weak acids, such as barbiturates, are usually much more soluble at higher pH values where they are ionized than at lower pH values where they are in the un-ionized acid form. Lowering the pH of their solutions well below their pKa may cause precipitation of the un-ionized acid. Depending on the supersaturation of the un-ionized acids or bases and on the presence of stabilizing agents, the resultant dispersions may be in the colloidal range.

Kinetics of Particle Formation—When the solubility of a compound in water is exceeded, its solution becomes supersaturated and the compound may precipitate or crystallize. The rate of precipitation, the particle size (whether colloidal or coarse), and the particle size uniformity or distribution (whether a narrow distribution and nearly monodisperse or homodisperse particles, or a broad distribution and polydisperse or heterodisperse particles) depend on two successive and largely independent processes, nucleation and growth of nuclei.

When a solution of a salt or of sucrose is supercooled, or when a chemical reaction produces a salt in a concentration exceeding its solubility product, separation of the excess solid from the supersaturated solution is far from instantaneous. Clusters of ions or molecules called nuclei must exceed a critical size before they become stable and capable of growing into colloidal size crystals. These embryonic particles have much more surface for a given weight of material than large and stable crystals, resulting in higher surface free energy and greater solubility.

Whether *nucleation* takes place depends on the *relative supersaturation*. If C is the actual concentration of the solute before crystallization has set in, and C_s is its solubility limit, $C - C_s$ is the supersaturation and $(C - C_s)/C_s$ is the relative supersaturation. Von Weimarn recognized that the rate or velocity of nucleation (number of nuclei formed per liter per second) is proportional to the relative supersaturation. Nucleation seldom occurs at relative supersaturations below 3. The foregoing statement refers to homogeneous nucleation, where the nuclei are clusters of the same chemical composition as the crystallizing phase. If the solution contains solid impurities, such as dust particles in suspension, these may act as nuclei or centers of crystallization (heterogeneous nucleation).

Once nuclei have formed, the second process, *crystallization*, begins. Nuclei grow by accretion of ions or molecules from solution forming colloidal or coarser particles until the supersaturation is relieved, ie, until $C = C_s$. The rate of

crystallization or growth of nuclei is proportional to the supersaturation. The appropriate equation,

$$\frac{dm}{dt} = \frac{A_{sp}D}{\delta}(C - C_s)$$

is similar to the Noyes-Whitney equation governing the dissolution of particles (see Chapter 31) except that $C < C_s$ for the latter process, making dm/dt negative. In both equations, m is the mass of material crystallizing out in time t, D is the diffusion coefficient of the molecules or ions of the solute, δ is the length of the diffusion path or the thickness of the liquid layer adhering to the growing particles, and A_{sp} is their specific surface area. The presence of dissolved impurities may affect the rate of crystallization and even change the crystal habit, provided that these impurities are surface-active and become adsorbed on the nuclei or growing crystals.[22,23,25-28] For instance, 0.005% polysorbate 80 or octoxynol 9 significantly retard the growth of methylprednisolone crystals in aqueous media. Gelatin or povidone, at concentrations <0.10%, retard the crystal growth of sulfathiazole in water.

Von Weimarn found that the particle size of the crystals depends strongly on the concentration of the precipitating substance. At a very low concentration and slight relative supersaturation, diffusion is quite slow because the concentration gradient is very small. Sufficient nuclei will usually form to relieve the slight supersaturation locally. Crystal growth is limited by the small amount of excess dissolved material available to each particle. Hence, the particles cannot grow beyond colloidal dimensions. This condition is represented by points A, D and G of the schematic plot of von Weimarn (Fig 19-20). At intermediate concentrations, the extent of nucleation is somewhat greater but much more material is available for crystal growth. Coarse crystals rather than colloidal particles result (points B, E or H).

At high concentrations, nuclei appear so quickly and in such large numbers that supersaturation is relieved almost immediately, before appreciable diffusion occurs. The high viscosity of the medium also slows down diffusion of excess dissolved ions or molecules, retarding crystal growth without substantially affecting the rate of nucleation. A large number of very small particles results which, because of their proximity, tend to link, producing a translucent gel (points C and F). On subsequent dilution with water, such gels usually yield colloidal dispersions.

Thus, colloidal systems are usually produced at very low and high supersaturations. Intermediate values of supersaturation tend to produce coarse crystals. Low solubility is a necessary condition for producing colloidal dispersions. If the solubility of the precipitate is increased, for instance by heating the dispersion, a new family of curves will result, similar in shape to ABC, DEF, and GHI of Fig 19-20, but displaced upwards (towards larger particle sizes) and to the right (towards higher concentrations).[25-28]

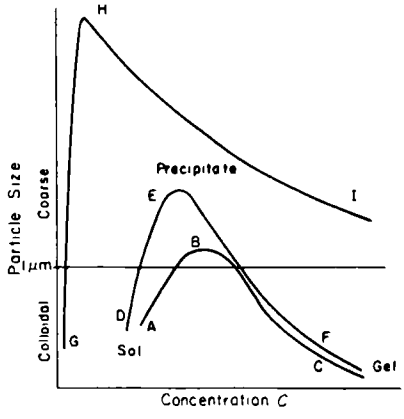


Fig 19-20. Effect of the concentration of the precipitating material and of aging on particle size.[28] Curves ABC, DEF and GHI correspond to increasing aging. Both axes are on a logarithmic scale.

Condensation methods generally produce polydisperse sols because nucleation continues while established nuclei grow. The particles in the resultant dispersion grew from nuclei formed at different times and had different growth periods.

A useful technique for preparing monodispersed sols in the colloidal range by precipitation consists in forming all the nuclei in a single, brief burst: When, in the course of the precipitation process, the rate of homogeneous nucleation becomes appreciable, a brief period of nucleation relieves the supersaturation partially to such an extent that no new nuclei form subsequently. By controlling the precipitation process, it is rendered so slow that the supersaturation remains too small for further nucleation. Therefore, the nuclei formed in the initial burst grow uniformly by diffusion of the precipitating material as the precipitation process proceeds slowly. Throughout the rest of the precipitation, the supersaturation never again reaches sufficiently high values for forming new nuclei. It is relieved by continuous growth of the existing nuclei.[23,25,31]

Controlled hydrolysis of salts of di- and trivalent cations in aqueous solution at elevated temperatures has been used to produce colloidal dispersions of metal (hydrous) oxides of uniform size and shape, in a variety of well-defined shapes (eg, sphere, lath, cube, disc, hexagonal). Complexation of the cations, concentration and temperature control the rate of hydrolysis and, hence, the chemical composition, crystallinity, shape and size of the dispersed phase.[32]

A feature of Fig 19-20 is that aging increases the particle size. Curves ABC, DEF and GHI correspond to increasing times after mixing the reagents. Typical ages are 10–30 min, several hours, and weeks or years, respectively. This gradual increase in particle size of crystals in their mother liquor is a recrystallization process called *Ostwald ripening*. Very small particles have a higher solubility than large particles of the same substance owing to their greater specific surface area and higher surface free energy. In a saturated solution containing precipitated particles of the solute in a wide range of particle sizes, the very smallest particles dissolve spontaneously and the material deposits onto the large particles. The growth of the large crystals at the expense of the very small ones occurs because this process lowers the free energy of the dispersion. As mentioned above, the most stable system is the suspension of a few coarse crystals, whereas the colloidal dispersion of a great many fine particles of the same substance is intrinsically less stable.

The spontaneous coarsening of colloidal dispersions on aging is accelerated by a relatively high solubility of the precipitate and can be retarded by lowering the solubility or by adding traces of surface-active compounds which are adsorbed at the particle surface. For instance, barium sulfate precipitated by mixing concentrated solutions of sodium sulfate and barium chloride is largely in the colloidal range and passes through filter paper. The colloidal particles gradually grow in size by Ostwald ripening, forming large crystals which can be removed quantitatively by filtration. Heating the aqueous dispersion speeds up this recrystallization by increasing the solubility of barium sulfate in water. The addition of ethyl alcohol lowers the solubility, retarding Ostwald ripening so that the dispersion remains in the colloidal state for years.

Mathematically the effect of particle size on solubility is expressed as

$$S = S_\infty \exp\left(\frac{2\gamma M}{r\rho RT}\right) \qquad (34)$$

Table XIII—Effects of Particle Size on Solubility

r (μm)	S
0.01	$7\,S_\infty$
0.10	$1.12\,S_\infty$
1.0	$1.01\,S_\infty$
10	$1.001\,S_\infty$

$M = 500$; $\gamma = 30$ ergs/cm^2; $\rho = 1$

where S is the solubility of a spherical crystal of radius r, S_∞ is the solubility of an infinitely large crystal ($r = \infty$), M is the molecular weight, ρ is the density, γ is the crystal/solvent interfacial tension, R is the gas constant and T is the absolute temperature. Only approximations can be obtained with this equation because the particles are not spheres, and γ values are different for different crystal faces. Table XIII shows the magnitude of particle size effects on the solubility for reasonable values of M, γ and ρ. It is evident that with particles in the colloidal range, ie, $r \geq 1$ μm, S values become appreciably greater than that for a coarse crystal, hence the tendency for very fine particles to dissolve and for coarse crystals to grow at the expense of the former. This difference in solubility explains why difficulty is encountered in preparing and stabilizing suspensions of very fine particles of certain substances.

Two techniques are used to increase the solubility of very slightly soluble drugs and, hence, their rate of dissolution *in vivo*. Many organic compounds exist in various polymorphic modifications. For instance, corticosterone, testosterone, sulfaguanidine and pentobarbital each have four polymorphic forms, with different melting points and crystal structures. The three metastable polymorphs have higher solubilities than the stable form. Solvates of solid drugs, eg, hydrates, have different crystalline structures and either higher or lower solubilities than the anhydrous forms. Theophylline monohydrate is less soluble than the anhydrous form while succinylsulfathiazole is less soluble than its solvate with 1-pentanol. Milling and grinding organic crystals may produce significant proportions of amorphous or strained crystalline material, which has higher solubility than the original crystalline material.[33]

Another process by which particles in colloidal dispersions grow in size is by agglomeration of individual particles into aggregates. This process, called coagulation, is discussed below.

Purification of Hydrosols by Dialysis and Ultrafiltration

Many hydrosols contain low molecular-weight, water-soluble impurities. Inorganic dispersions often contain salts formed by the reaction producing the disperse phase. Salts are especially objectionable in the case of hydrophobic dispersions because they tend to coagulate such dispersions. Protein solutions often contain salts added as part of the separation procedure. The blood of patients with renal insufficiency contains excessive concentrations of urea and other low-molecular-weight metabolites and salts. These dissolved impurities of small molecular size are removed from the colloidal dispersions by means of membranes with pore openings smaller than the colloidal particles.

Membranes—Conventional filter papers are permeable to colloidal particles as well as to small solute molecules. Among the early membranes capable of retaining colloidal particles but permeable to small solute molecules were pig's bladder and parchment. Most membranes in current use consist of cellulose, cellulose nitrate prepared from collodion, cellulose acetate or synthetic polymers, and are available in a variety of shapes, gauges, and pore sizes. *Gel cellophane* is most widely used. It consists of sheets or tubes of cellulose made by extruding cellulose xanthate solutions (viscose) through slit or annular dies into a sodium bisulfate/sulfuric acid bath which decomposes the xanthate, precipitating the regenerated cellulose in a highly swollen or gel state. If the cellulose film were permitted to dry after purification and washing with water, it would crystallize and shrink excessively, losing most of its extensive micropore structure and turning somewhat brittle. The film is therefore impregnated with glycerin before drying. Glycerin remains in the film rather than evaporating like water. It reduces the shrinkage and blocks crystallization. This action prevents the collapse of the porous gel structure and plasticizes the film, keeping it flexible. A typical dialysis tube made from sausage casing swells to about twice its thickness in water and has an average pore diameter of 34 Å. While the pore structure of cellophane films used in dialysis and ultrafiltration causes retention of colloidal particles but permits the passage of small solute molecules, osmotic membranes are only permeable to water and retain small solute molecules as well as colloidal particles.

Dialysis—The colloidal dispersion is placed inside a sac made of sausage casing dipping in water. The small solute molecules diffuse out into the water while the colloidal material remains trapped inside because of its size. The rate of dialysis is increased by increasing the area of the membrane, by stirring, and by maintaining a high concentration gradient across the membrane. For the latter purpose, the water is replenished continuously or at least frequently. A membrane configuration which provides a particularly extensive transfer area for a given volume of dispersion is the hollow fiber. A typical fiber measures 175 μm inside diameter and 225 μm outside diameter. The dispersion to be dialyzed is circulated inside a bundle of parallel fibers while water is circulated outside the fibers throughout the bundle. Dialysis of the diffusing species takes place across the thin fiber wall. Dialysis is used in the laboratory to purify sols and to study binding of drugs by proteins, as well as in some manufacturing processes.

Electrodialysis—If the low-molecular-weight impurities to be removed are electrolytes, the dialysis can be speeded up by applying an electric potential to the sol which produces electrolysis. An electrodialyzer (Fig 19-21) is divided into three compartments by two dialysis membranes supported by screens. The two outer compartments, in which the two electrodes are placed, are filled with water while the sol is placed into the center compartment. Under the influence of the applied potential, the anions migrate from the sol into the anode (right) compartment while the cations migrate into the cathode compartment. Low-molecular-weight nonelectrolyte solutes diffuse into either compartment.

Colloidal particles are usually charged and therefore tend to migrate towards the membrane sealing off the compartment with the electrode of opposite charge. The combination of electrophoresis (see below) and gravitational sedimentation produces the accumulation of negatively charged sol particles shown in Fig 19-21. Hence the supernatant

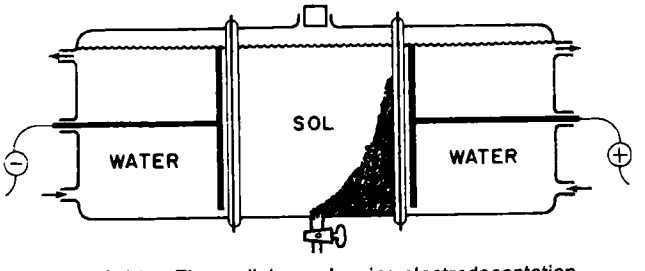


Fig 19-21. Electrodialyzer showing electrodecantation

liquid can be changed by decantation. This process, which may be used to speed up electrodialysis, is called *electrodecantation.*[21,25]

Ultrafiltration—When a sol is placed in a compartment closed by a dialysis membrane and pressure is applied, the liquid and the small solute molecules are forced through the membrane while the colloidal particles are retained. This process, called ultrafiltration, is based on a sieving mechanism in which all components smaller than the pore size of the filter membrane pass through it. The pressure difference required to push the dispersion medium through the ultrafilter is provided by gas pressure applied on the sol side or by suction on the filtrate side. The membrane is usually supported on a fine wire screen.[24-27]

As ultrafiltrate is being removed, the sol becomes more concentrated because a constant amount of disperse particles is confined to a decreasing volume of liquid. Some dissolved small molecules or ions are left in the sol together with the residual water. To avoid the increase in concentration of the colloidal particles and remove the dissolved impurities completely, the ultrafiltrate squeezed from the sol is replenished continuously or intermittently with an equal volume of water. During ultrafiltration, solids tend to accumulate on and near the membrane. To prevent this buildup and maintain uniform composition throughout the sol, it is stirred.

Bundles of hollow fibers are used for ultrafiltration in the laboratory and on large scale. To withstand higher pressures, the wall thickness of the fibers used in ultrafiltration is usually greater than that of fibers used exclusively for dialysis. When hollow fibers are fouled by excessive accumulation of solids on the inner wall, they are cleaned by backflushing with water or ultrafiltrate.

Hemodialysis—The blood of uremic patients is dialyzed periodically in "artificial kidney" dialyzers to remove urea, creatinine, uric acid, phosphate and other metabolites, and excess sodium and potassium chloride. The dialyzing fluid contains sodium, potassium, calcium, chloride and acetate ions (the latter are converted in the body to bicarbonate), dextrose and other constituents in the same concentration as normal plasma. Since it contains no urea, creatinine, uric acid, phosphate nor any of the other metabolites normally eliminated by the kidneys, these compounds diffuse from the patient's blood into the dialyzing fluid until their concentration is the same in blood and fluid. Sodium and potassium chloride diffuse from blood to fluid because of their higher initial concentration in the blood, and continue to diffuse until the concentration is equalized. The volume of dialyzing fluid is much greater than that of blood. The great disparity in volume and the replenishment of dialyzate with fresh fluid ensure that the metabolites and the excess of electrolytes are removed almost completely from the blood. Hemodialysis is also employed in acute poisoning cases.

Plasma proteins and blood cells cannot pass through the dialysis membrane because of their size. Edema resulting from water retention can be relieved by ultrafiltration through the application of a slight pressure on the blood side or a partial vacuum on the fluid side.

The three geometries used to circulate the blood and the dialyzing fluid in a countercurrent fashion are a coil of flattened cellulose tubing wound concentrically with a supporting mesh screen around a core, a stack of flat cellulose sheets separated by ridged or grooved plates, and hollow fibers. The regenerated cellulose used in the former two is precipitated from a cuprammonium solution. The hollow cellulose acetate fibers have an outside diameter of about 270 μm and a wall thickness of 30 μm.[34] The advantage of hollow fibers is their compactness. A bundle of 10,000 fibers 18 cm long has a surface area of 1.4 m^2.

Particle Shape, Optical, and Transport Properties of Lyophobic Dispersions

Hydrophobic materials handled by pharmacists in aqueous dispersion range from metallic conductors to inorganic precipitates to organic solids and liquids which are electric insulators. Despite the great diversity of the hydrophobic disperse phase, their hydrosols have certain common characteristics.

Particle Shape and Particle Size Distribution—Both of these properties depend on the chemical and physical nature of the disperse phase and on the method employed to prepare the dispersion. Primary particles exist in a great variety of shapes. Their aggregation produces an even greater variety of shapes and structures. Precipitation and mechanical comminution generally produce randomly shaped particles unless the precipitating solids possess pronounced crystallization habits or the solids being ground possess strongly developed cleavage planes. Precipitated aluminum hydroxide gels and micronized particles of sulfonamides and other organic powders have typical irregular random shapes. An exception is bismuth subnitrate. Even though its particles are precipitated by hydrolyzing bismuth nitrate solutions with sodium carbonate, its particles are lath-shaped. Precipitated silver chloride particles have a cubic habit which is apparent under the electron microscope. Lamellar or plate-like solids in which the molecular cohesion between layers is much weaker than within layers frequently preserve their lamellar shape during mechanical comminution, because milling and micronization break up stacks of thin plates in addition to fragmenting plates in the lateral dimensions. Examples are graphite, mica and kaolin. Figure 19-22 shows a Georgia crude clay as mined. Processing yields the refined, fine-particle kaolinite of Fig 19-23. Similarly, macroscopic asbestos and cellulose fibers consist of bundles of microscopic and submicroscopic fibrils. Mechanical comminution or beating splits these bundles into the component fibrils of very small diameters as well as cutting them shorter.

Microcrystalline cellulose is a fibrous thickening agent and tablet additive made by selective hydrolysis of cellulose.

Fig 19-22. Scanning electron micrograph of a crude kaolin clay as mined. Processing yields the fine particle material of Fig 19-23 (courtesy, John L Brown, Engineering Experiment Station, Georgia Institute of Technology).

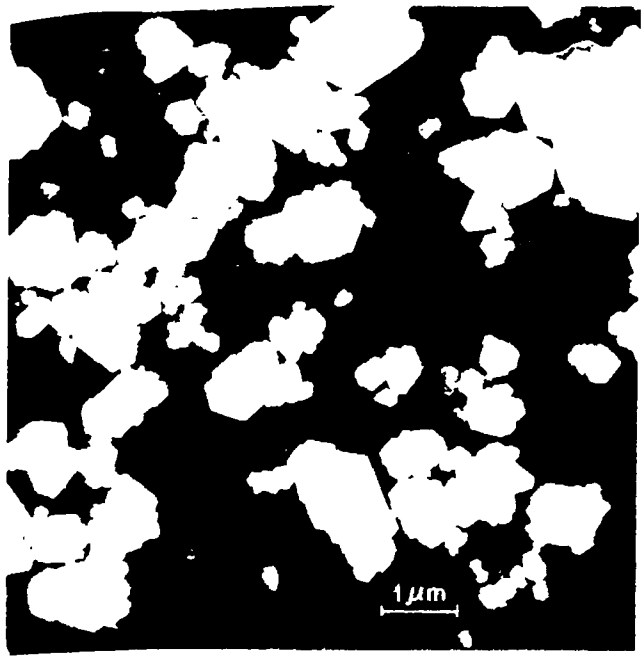


Fig 19-23. Transmission electron micrograph of a well crystallized, fine-particle kaolin. Note hexagonal shape of the clay platelets (courtesy, John L Brown, Engineering Experiment Station, Georgia Institute of Technology).

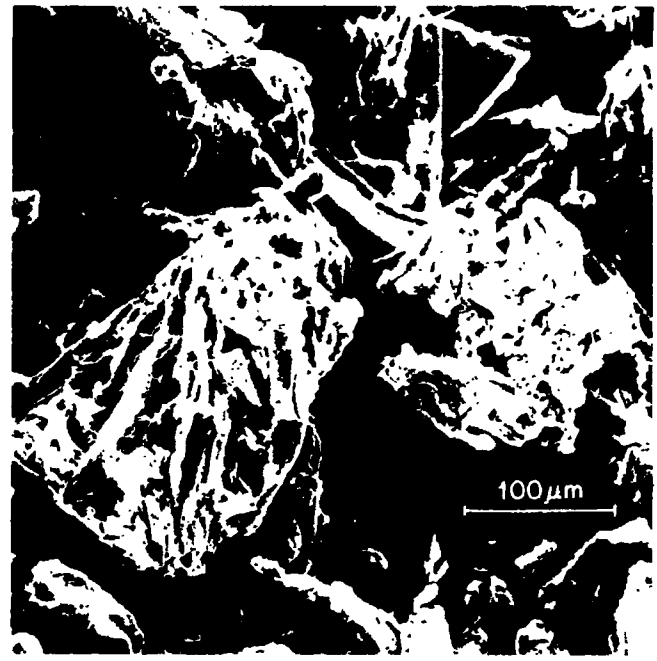


Fig 19-24. Scanning electron micrograph of Avicel PH-102 tableting grade microcrystalline cellulose. The aggregates of fiber bundles are porous and compressible (courtesy, FMC Corporation; Avicel is a registered trademark of FMC Corporation).

Native cellulose consists of crystalline regions where the polymer chains are well aligned and in registry, with maximum interchain attraction by secondary valence forces, called crystallites, and of more disordered regions having lower density and reduced interchain attraction and crystallinity, the so-called "amorphous" regions. During treatment with dilute mineral acid, the acid penetrates the amorphous regions relatively fast and hydrolyzes the polymer chains into water-soluble fragments. If the acid is washed out before it penetrates the crystalline regions appreciably, the crystallites remain intact. Wet milling and spray-drying the aqueous suspension produces spongy and porous aggregates of rod-shaped or fibrillar bundles shown in Fig 19-24. These aggregates, averaging 100 μm in size, were embrittled by the acid treatment and lost the elasticity of the native cellulose. They are well compressible and capable of undergoing plastic deformation, a property important in tableting. Their porosity permits the aggregates to absorb liquid ingredients while still remaining a free-flowing powder, thus preventing these liquids from reducing the flowability of the granulation or direct-compression mass during tableting. The swelling of the cellulosic particles in water speeds up the disintegration of the ingested tablets.

Additional shear breaks up the aggregated bundles into the individual, needle- or rod-shaped cellulose crystallites shown in Fig 19-25. The latter, which average 0.3 μm in length and 0.02 μm in width, are of colloidal dimensions. These primary particles act as suspending agents in water, producing thixotropic structured vehicles. At concentrations above 10%, eg 14 or 15%, the cellulose microcrystals gel water to ointment consistency by swelling and producing a continuous network of rods extending throughout the entire vehicle. Attraction between the elongated particles is presumably due to flocculation in the secondary minimum (see below). Treatment of the microcrystalline mass with sodium carboxymethylcellulose facilitates its disintegration into the primary needle-shaped particles and enhances their thickening action.

While in the special cases of certain clays and cellulose, comminution produces lamellar and fibrillar particles, respectively, as a rule regular particle shapes are produced by

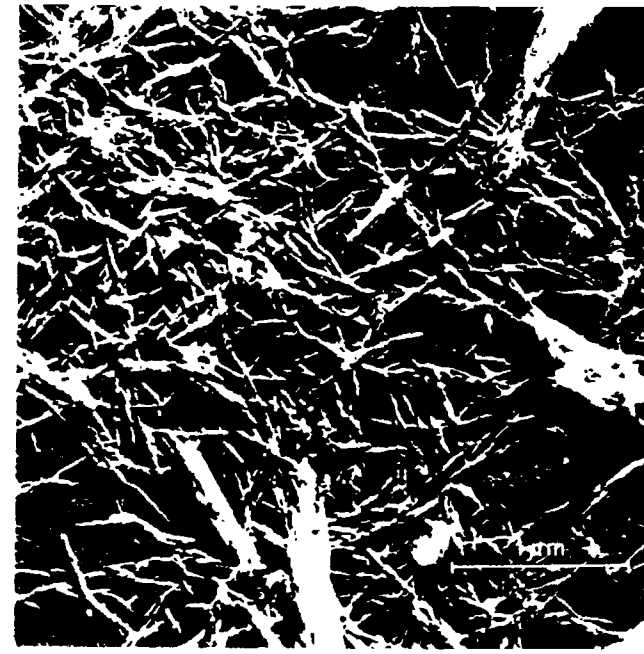


Fig 19-25. Transmission electron micrograph of Avicel RC-591 thickening grade microcrystalline cellulose. The needles are individual cellulose crystallites; some are aggregated into bundles (courtesy, FMC Corporation; Avicel is a registered trademark of FMC Corporation).

condensation rather than by disintegration methods. *Colloidal silicon dioxide* is called fumed or pyrogenic silica because it is manufactured by high-temperature, vapor-phase hydrolysis of silicon tetrachloride in an oxy-hydrogen flame, ie, a flame produced by burning hydrogen in a stream of oxygen. The resultant white powder consists of submicroscopic spherical particles of rather uniform size (narrow particle size distribution). Different grades are produced by different reaction conditions. Relatively large, single

spherical particles are shown in Fig 19-26. Their average diameter is 50 nm (500 Å), corresponding to the comparatively small specific surface area of 50 m^2/g. Smaller spherical particles have correspondingly larger specific surface areas; the grade with the smallest average diameter, 5 nm, has a specific surface area of 380 m^2/g. During the manufacturing process, the finer-grade particles tend to sinter or grow together into chain-like aggregates resembling pearl necklaces or streptococci (see Fig 19-27).

Since fumed silica is amorphous, its inhaled dust causes no silicosis. The spheres of colloidal silicon dioxide are nonporous. While the density of the spherical particles is 2.13 g/cm^3, the bulk density of their powder is a mere 0.05 g/cm^3; the powder is extremely light. This results in two pharmaceutical and cosmetic applications for colloidal silicon dioxide. It is used to increase the fluffiness or bulk volume of powders. Even more than microcrystalline cellulose, the high porosity of silica enables it to absorb a variety of liquids from fluid fragrances to viscous tars, transforming them into free-flowing powders that can be incorporated into tablets or capsules. The porosity in colloidal silicon dioxide is due entirely to the enormous void space between the particles, which themselves are solid.

When these ultrafine particles are incorporated at levels as low as 0.1 to 0.5% into a powder consisting of coarse particles or granules, they coat the surface of the latter and act as tiny ball bearings and spacers, improving the flowability of the powder and eliminating caking. This action is important in tableting. Moreover, colloidal silicon dioxide improves tablet disintegration.

The surface of the particles contains siloxane (Si—O—Si) and silanol (Si—OH) groups. When colloidal silicon dioxide powder is dispersed in nonpolar liquids, the particles tend to adhere to one another by hydrogen bonds between their surface groups. With finer grades of colloidal silicon dioxide, the spherical particles are linked together into short chain-like aggregates as shown in Fig 19-27, thus agglomerating into loose three-dimensional networks which increase the viscosity of the liquid vehicles very effectively at levels as low as a few percent. These hydrogen-bonded structures are torn apart by stirring but rebuilt while at rest, conferring thixotropy to the thickened liquids.

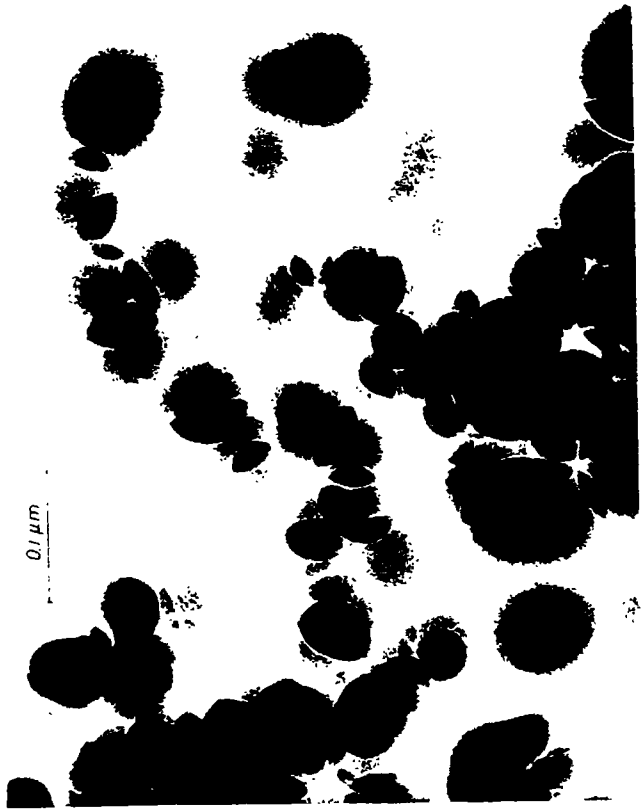


Fig 19-26. Transmission electron micrograph of Aerosil OX 50, ground and dusted on. The spheres are translucent to the electron beam, causing overlapping portions to be darker owing to increased thickness (courtesy, Degussa AG of Hanau, West Germany; Aerosil is a registered trademark of Degussa). The suffix 50 indicates the specific surface area in m^2/g.

Fig 19-27. Transmission electron micrograph of Aerosil 130, ground and dusted on. The spheres are fused together into chain-like aggregates (courtesy, Degussa AG of Hanau, West Germany; Aerosil is a registered trademark of Degussa). The suffix 130 gives the specific surface area in m^2/g.

The grades which consist of relatively large and unattached spherical particles, such as those of Fig 19-26, are less efficient thickening agents as they lack the high specific surface area and the asymmetry of the finer grades, which consist of short chains of fused spherical particles. In the latter category is Aerosil 200, the grade most widely used as a pharmaceutical adjuvant, whose primary spheres, which are extensively sintered together, have an average diameter of 12 nm. At levels of 8 to 10%, it thickens liquids of low polarity such as vegetable and mineral oils to the consistency of ointments, imparting considerable yield values to them. The consistency of ointments thickened with colloidal silicon dioxide is not appreciably reduced at higher temperatures. Incorporation of colloidal silicon dioxide into ointments and pastes, such as those of zinc oxide, also reduces the syneresis or *bleeding* of the liquid vehicles.

Hydrogen-bonding liquids like alcohols and water solvate the silica spheres, reducing the hydrogen bonding between particles. These solvents are gelled at silica levels of 12–18% or higher.

Latexes of polymers are aqueous dispersions prepared by emulsion polymerization. Their particles are spherical because polymerization of solubilized liquid monomer takes

place inside spherical surfactant micelles which swell because additional monomer keeps diffusing into the micelles. Examples include latex-based paints. Some *clays* grow as plate-like particles possessing straight edges and hexagonal angles, eg bentonite and kaolin (see Fig 19-23). Other clays have lath-shaped (nontronite) or needle-shaped particles (attapulgite).

Emulsification produces spherical droplets to minimize the oil-water interfacial area. Cooling the emulsion below the melting point of the disperse phase freezes it in the spherical shape. For instance, paraffin can be emulsified in 80° water; cooling to room temperature produces a hydrosol with spherical particles.

Sols of viruses and globular proteins, which are hydrophilic, contain compact particles possessing definite geometric shapes. Poliomyelitis virus is spherical, tobacco mosaic virus is rod-shaped, while serum albumin and the serum globulins are prolate ellipsoids of revolution (football-shaped).

Dispersion methods produce sols with wide particle size distributions. Condensation methods *may* produce essentially monodisperse sols provided specialized techniques are employed. Monodisperse polystyrene latexes are available for calibration of electron micrographs (see Fig 19-23). Biologic hydrophilic polymers, such as nucleic acids and proteins, form largely monodisperse particles, as do more highly organized structures such as lipoproteins and viruses.

Light-Scattering by Colloidal Particles—The optical properties of a medium are determined by its refractive index. When the refractive index is uniform throughout, light will pass the medium undeflected. Whenever there are discrete variations in the refractive index caused by the presence of particles or by small-scale density fluctuations, part of the light will be scattered in all directions. An optical property characteristic of colloidal systems, called the *Tyndall beam*, is familiar to everyone in the case of aerosols. When a narrow beam of sunlight is admitted through a small hole into a darkened room, the presence of the minute dust particles suspended in air is revealed by bright flashing points.

A beam of light striking a particle polarizes the atoms and molecules of that particle, inducing dipoles which act as secondary sources and reemit weak light of the same wavelength as the incident light. This phenomenon is called *light-scattering*. The scattered radiation propagates in all directions away from the particle. In a bright room, the light scattered by the dust particles is too weak to be noticeable.

Colloidal particles suspended in a liquid also scatter light. When an intense, narrowly defined beam of light is passed through a suspension, its path becomes visible because of the scattering of light by the particles in the beam. This Tyndall beam becomes most visible when viewed against a dark background in a direction perpendicular to the incident beam. The magnitude of the turbidity or opalescence depends on the nature, size and concentration of the particles. When clear mineral oil is dispersed in an equal volume of a clear aqueous surfactant solution, the resultant emulsion is milky white and opaque due to light scattering. Microemulsions, where the emulsified droplets are about 40 nm (400 Å) in diameter, ie, much smaller than the wavelength of visible light, are transparent and clear to the naked eye.

The *dark-field microscope* or *ultramicroscope*, which permits observation of particles much smaller than the wavelength of light, was the only means of detecting submicroscopic particles before the advent of electron microscopy. A special cardioid condenser produces a hollow cylinder of light and converges it into a hollow cone focused on the sample. The sample is at the apex of the cone, where the light intensity is high. After passing through the sample, the cone of light diverges and passes outside of the microscope objective. A homogeneous sample thus gives a dark field. A similar effect can be produced with a regular Abbe condenser outfitted with a central stop and a strong light source. Colloidal particles scatter light in all directions. Some of the scattered light enters the objective and shows up the particles as bright spots. Thus, even particles smaller than the wavelength of light can be detected, provided their refractive index differs sufficiently from that of the medium. Dissolved polymer molecules and highly solvated gel particles do not scatter enough light to become visible. Asymmetric particles like flat bentonite platelets give flashing effects as they rotate in Brownian motion, because they scatter more light with their basal plane perpendicular to the light beam than edgewise. Brownian motion, sedimentation, electropboretic mobility, and the progress of flocculation can be studied with the dark-field microscope. Polydispersity can be estimated qualitatively because larger particles scatter more light and appear brighter. The resolving power of the ultramicroscope is no greater than that of the ordinary light microscope. Particles closer together than 0.2 μm appear as a single blur.

Turbidity may be used to measure the concentration of dispersed particles in two ways. In *turbidimetry*, a spectrophotometer or photoelectric colorimeter is used to measure the intensity of the light transmitted in the incident direction. Turbidity, τ, is defined by an equation analogous to Beer's law for the absorption of light (see Chapter 30),[24,25,27] namely

$$\tau = \frac{1}{l} \ln \frac{I_0}{I_t}$$

where I_0 and I_t are the intensities of the incident and transmitted light beams, and l is the length of the dispersion through which the light passes.

If the dispersion is less turbid, the intensity of light scattered at 90° to the incident beam is measured with a *nephelometer*. Both methods require careful standardization with suspensions containing known amounts of particles similar to those to be measured. The concentration of colloidal dispersions of inorganic and organic compounds and of bacterial suspensions can thus be measured by their turbidity.

The turbidity or Tyndall effect of hydrophilic colloidal systems like aqueous solutions of gums, proteins and other polymers is far weaker than that of lyophobic dispersions. These solutions appear clear to the naked eye. Their turbidity can be measured with a photoelectric cell/photomultiplier tube and serves to determine the molecular weight of the solute.

The theory of light scattering was developed in detail by Lord Rayleigh. For white nonabsorbing nonconductors or dielectrics like sulfur and insoluble organic compounds, the equation obtained for spherical particles whose radius is small compared to the wavelength of light λ is[24-27]

$$I_s = I_0 \frac{4\pi^2 n_0^2 (n_1 - n_0)^2}{\lambda^4 d^2 c} (1 + \cos^2 \theta)$$

I_0 is the intensity of the unpolarized incident light; I_s is the intensity of light scattered in a direction making an angle θ witb the incident beam and measured at a distance d. The scattered light is largely polarized. The concentration c is expressed as the number of particles per unit volume. The refractive indices n_1 and n_0 refer to the dispersion and the solvent, respectively.

Since the intensity of scattered light is inversely proportional to the fourth power of the wavelength, blue light ($\lambda \cong$ 450 nm or 4500 Å) is scattered much more strongly than red light ($\lambda \cong$ 650 nm or 6500 Å). With incident white light, colloidal dispersions of colorless particles appear blue when

viewed in scattered light, ie, in lateral directions such as 90° to the incident beam. Loss of the blue rays due to preferential scattering leaves the transmitted light yellow or red. Preferential scattering of blue radiation sideways accounts for the blue color of the sky, sea, cigarette smoke, and diluted milk and for the yellow-red color of the rising and setting sun viewed head-on.

The particles in pharmaceutical suspensions, emulsions and lotions are generally larger than the wavelength of light λ. When the particle size exceeds $\lambda/20$, destructive interference between light scattered by different portions of the same particle lowers the intensity of scattered light and changes its angular dependence. Rayleigh's theory was extended to large and to strongly absorbing and conducting particles by Mie and to nonspherical particles by Gans.[21,22,24-27] By using appropriate precautions in experimental techniques and in interpretation, it is possible to determine an average particle size and even the particle size distribution of colloidal dispersions and coarser suspensions by means of turbidity measurements.

Diffusion and Sedimentation—The molecules of a gas or liquid are engaged in a perpetual, random thermal motion which causes them to collide with one another and with the container wall billions of times per second. Each collision changes the direction and the velocity of the molecules involved. Dissolved molecules and suspended colloidal particles are continuously and randomly buffeted by the molecules of the suspending medium. This random bombardment imparts to solutes and particles an equally unceasing and erratic movement called *Brownian motion*, after the botanist Robert Brown who first observed it under the microscope with an aqueous pollen suspension. The Brownian motion of colloidal particles mirrors on a magnified scale the random movement of the molecules of the liquid or gaseous suspending medium, and represents a three-dimensional random walk.

Solute molecules and suspended colloidal particles undergo rotational and translational Brownian movement. For the latter, Einstein derived the equation

$$\bar{x} = \sqrt{2Dt}$$

where $\bar{x}$ is the mean displacement in the x-direction in time t and D is the *diffusion coefficient*. Einstein also showed that for spherical particles of radius r under conditions specified in Chapter 20 for the validity of Stokes' law and Einstein's law of viscosity

$$D = \frac{RT}{6\pi\eta rN}$$

where R is the gas constant, T the absolute temperature, N Avogadro's number, and η the viscosity of the suspending medium.

The diffusion coefficient is a measure of the mobility of a dissolved molecule or suspended particle in a liquid medium. Representative values at room temperature, in cm^2/sec, are 4.7×10^{-6} for sucrose and 6.1×10^{-7} for serum albumin in water. With a diffusion coefficient of 1×10^{-7} cm^2/sec, Brownian motion causes a particle to move by an average distance of 1 cm in one direction in 58 days, by 1 mm in 14 hr, and by 1 μm in 0.05 sec. Smaller molecules diffuse faster in a given medium. Assuming spherical shape, the radius of a serum albumin molecule is 35 Å and that of a sucrose molecule 4.4 Å. The ratio of the radii of the two molecules 35/4.4 = 7.9, is nearly identical with the inverse ratio of their diffusion coefficients in water, $4.7 \times 10^{-6}/6.1 \times 10^{-7} = 7.7$, in agreement with the above equation. Diffusion coefficients of steroids and other molecules of similar size dissolved in absorption bases based on petrolatum are generally in the 10^{-10} to 10^{-8} cm^2/sec range. Steroids have only slightly higher molecular weights than sucrose. Their much smaller diffusion coefficients are due to the much higher viscosity of the vehicle.

Dynamic light-scattering or photon-correlation spectroscopy is based on the fact that the light scattered by particles in Brownian motion undergoes a minute shift in wavelength by the usual Doppler effect. The shift is so small that it can be detected only by laser light beams, which are strictly monochromatic and very intense. The wavelength shift, which shows up as line broadening, is used to determine the diffusion coefficient of the particles,[23,26] which in turn yields their radius according to the equation above.

Brownian motion and convection currents maintain dissolved molecules and small colloidal particles in suspension indefinitely. As the particle size and r increase, the Brownian motion decreases; $\bar{x}$ is proportional to $r^{-1/2}$. Provided that the density of the particle d_P and of the liquid vehicle d_L are sufficiently different, larger particles have a greater tendency to settle out when $d_P > d_L$ or to rise to the top of the suspension when $d_P < d_L$ than smaller particles of the same material.

The rate of *sedimentation* is expressed by the Stokes' equation (Eq 35), which can be rewritten as

$$h = \frac{2(d_P - d_L)r^2gt}{9\eta}$$

where h is the height through which a spherical particle settles in time t. The rate of sedimentation is proportional to r^2. Thus, with increasing particle size, the Brownian motion diminishes while the tendency to sediment increases. The two become equal for a critical radius when the distance h through which the particle settles equals the mean displacement $\bar{x}$ due to Brownian motion in the same time interval t.[35] In most pharmaceutical suspensions, sedimentation prevails. Intravenous vegetable oil emulsions do not tend to cream because the mean droplet size, *ca* 0.5 μm, is smaller than the critical radius.

Passive diffusion caused by a concentration gradient and carried out through Brownian motion is important in the release of drugs from topical preparations (see Chapter 87) and in the gastrointestinal absorption of drugs (see Chapter 35).

Viscosity—Most lyophobic dispersions have viscosities not much greater than that of the liquid vehicle. This holds true even at comparatively high volume fractions of the disperse phase unless the particles form continuous network aggregates throughout the vehicle, in which case yield values are observed. Most O/W and W/O emulsions have specific viscosities not much greater than those predicted by Einstein's modified law of viscosity (see Eq 11 of Chapter 20 and text). For instance, emulsions containing 40% v/v of the internal phase generally have viscosities only three to five times higher than that of the continuous phase. By contrast, the apparent viscosities of lyophilic dispersions, especially of polymer solutions, are several orders of magnitude greater than the viscosity of the solvent or vehicle even at concentrations of only a few percent solids. Lyophilic dispersions are also generally much more pseudoplastic or shear-thinning than lyophobic dispersions (see Chapter 20).

Electric Properties and Stability of Lyophobic Dispersions

Difference between Lyophilic and Lyophobic Dispersions—*Lyophilic* or solvent-loving solids are called hydrophilic if the solvent is water. Owing to the presence of high concentrations of hydrophilic groups, they dissolve or disperse spontaneously in water as far as is possible without breaking covalent bonds. Among hydrophilic groups are ionized ones which dissociate into highly hydrated ions like carboxylate, sulfonate or alkylammonium ions, and organic

functional groups like hydroxyl, carbonyl, amino, and imino which bind water through hydrogen bonding.

The free energy of dissolution or dispersion, ΔG_s, of hydrophilic solids includes a large negative (exothermic) heat or enthalpy of solvation, ΔH_s, and a large increase in entropy, ΔS_s. Since $\Delta G_s = \Delta H_s - T\Delta S_s$, ΔG_s has a large negative value: the dissolution of hydrophilic macromolecules and the dispersion of hydrophilic particulate solids in water occur spontaneously (see Chapter 16), overcoming the parallel increases in surface area and surface free energy. Dissolution and dispersion take place so that water can come into contact and interact with the hydrophilic groups of the solids (enthalpy of solvation), and to increase the number of available configurations of the macromolecules and particles (entropy increase).

The van der Waals energies of attraction between dissolved macromolecules or dispersed hydrophilic solid particles are smaller than ΔG_s and are, therefore, insufficient to cause separation of a solid polymer phase or agglomeration through flocculation or coagulation of the dispersed particles. Furthermore, the hydration layer surrounding dissolved macromolecules and dispersed particles forms a barrier preventing their close approach.

Hydrophobic solids and liquids such as organic compounds consisting largely of hydrocarbon portions with few if any hydrophilic functional groups, like cholesterol and other steroids, and some nonionized inorganic substances like sulfur, are hydrated slightly or not at all. Hence they do not disperse or dissolve spontaneously in water: ΔG_s is positive because of a positive (endothermic) ΔH_s term, making the reverse process (agglomeration) the spontaneous one. Aqueous dispersions of such hydrophobic solids or liquids can be prepared by physical means which supply the appropriate energy to the system (see above). They are unstable, however. The van der Waals attractive forces between the particles cause them to aggregate, since the solvation forces which promote dispersal in water are weak. If aqueous dispersions of hydrophobic solids are to resist reaggregation (coagulation and flocculation). they must be stabilized. Stabilizing factors include electric charges at the particle surface (due to dissociation of ionogenic groups of the solid or pertaining to adsorbed ions such as ionic surfactants) and the presence of adsorbed macromolecules or nonionic surfactants. These stabilizing factors do not alter the intrinsic thermodynamic instability of lyophobic dispersions; ΔG_s is still positive so that the reverse process of phase separation or aggregation is energetically favored over dispersal. They establish kinetic barriers which delay the aggregation processes almost indefinitely; the dispersed particles cannot come together close enough for the van der Waals attractive forces to produce coagulation.[24,26,27] These stabilization mechanisms are discussed below.

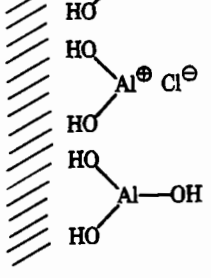
The reductions in surface area and surface free energy accompanying flocculation or coagulation are small because irregular solid particles, being rigid, touch only at a few points upon aggregation. The loose initial contacts may grow with time by sintering or recrystallization. Sintering consists of the "fusion" of primary particles into larger primary particles which propagates from initial small areas of contact. This recrystallization process is spontaneous because it decreases the specific surface area of the disperse solid and the surface free energy of the dispersion. Sintering is analogous to Ostwald ripening, the recrystallization process of transferring solid from colloidal to coarse particles discussed above. Low solubility and the presence of adsorbed surface-active substances retard both processes.

Origin of Electric Charges—Particles can acquire charges from several sources. In *proteins*, one end group of the polypeptide chain and aspartic and glutamic acid units contribute carboxylic acid groups, which are ionized into carboxylate ions in neutral to alkaline media. The other chain end group and lysine units contribute amino groups, arginine units contribute guanidine groups, and histidine units contribute imidazole groups. The nitrogen atoms of these groups become protonated in neutral to acid media. For electroneutrality, these cationic groups require anions, such as Cl^- if hydrochloric acid was used to make the medium acid and to supply the protons. The neutralizing ions, called counterions, dissociate from the ionogenic basic functional groups and can be replaced by other ions of like charge: they are not an integral part of the protein particle but are located in its immediate vicinity. The alkylammonium, guanidinium and imidazolium ions, which are attached to the protein molecule by covalent bonds, confer a positive charge to it. In neutral and alkaline media, Na^+, K^+, Ca^{2+} and Mg^{2+} are among the counterions neutralizing the negative charges of the carboxylate groups. The latter are covalently attached to and constitute an integral part of the protein particle, conferring a negative charge to it.

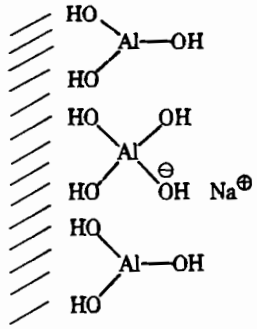
At an intermediate pH value, which ranges from 4.5 to 7 for the various proteins, the carboxylate anions and the alkylammonium, guanidinium, and imidazolium cations neutralize each other exactly. There is no need for counterions since the ionized functional groups which are an integral part of the protein molecule are in exact balance. At this pH value, called the *isoelectric point*, the protein particle or molecule is neutral; its electric charge is neither negative nor positive, but zero.[22,24,27]

Many other organic polymers contain ionic groups and are, therefore, called *polyelectrolytes* (polymeric electrolytes or salts). Natural polysaccharides of vegetable origin such as acacia, tragacanth, alginic acid and pectin contain carboxylic acid groups, which are ionized in neutral to alkaline media. Agar and carrageenan as well as the animal polysaccharides heparin and chondroitin sulfate, contain sulfuric acid hemiester groups, which are strongly acidic and ionize even in acid media. Cellulosic polyelectrolytes include *sodium carboxymethylcellulose*, while synthetic carboxylated polymers include *carbomer*, a copolymer of acrylic acid.

Aluminum hydroxide, $Al(OH)_3$, is dissolved by acids and alkalis forming aluminum ions, Al^{3+}, and aluminate ions, $[Al(OH)_4]^-$, respectively. In neutral or weakly acid media, at acid concentrations too low to cause dissolution, an aluminum hydroxide particle has some positive charges attributable to incompletely neutralized positive Al^{3+} valences. The portion of the surface of an aluminum hydroxide particle represented schematically below has one such positive charge neutralized by a Cl^- counterion:



In weakly alkaline media, at base concentrations too low to transform the aluminum hydroxide particles completely into aluminate and dissolve them, they bear some negative charges due to the presence of a few aluminate groups. The portion of the particle surface represented schematically below has one such negative group neutralized by a Na^+ counterion:



At a pH of 8.5 to 9.1,[36,37] there are neither $[Al(OH)_2]^+$ nor $[Al(OH)_4]^-$ ions in the particle surface but only neutral $Al(OH)_3$ molecules. The particles have zero charge and therefore need no counterions for charge neutralization. This pH is the isoelectric point. In the case of inorganic particulate compounds such as aluminum hydroxide, it is also called zero point of charge.

Bentonite clay is a lamellar aluminum silicate. Each lattice layer consists of a sheet of hydrated alumina sandwiched between two silica sheets. Isomorphous replacement of Al^{3+} by Mg^{2+} or of Si^{4+} by Al^{3+} confers net negative charges to the thin clay lamellas in the form of cation-exchange sites resembling silicate ions built into the lattice. The counterions producing electroneutrality are usually Na^+ (sodium bentonite) or Ca^{2+} (calcium bentonite). The zero point of charge is probably close to that of quartz, silica gel and other silicates, namely, at a pH of about 1.5 to 2.

Silver iodide sols can be prepared by the reaction

$$AgNO_3 + NaI \rightarrow AgI(s) + NaNO_3$$

In the bulk of the silver iodide particles, there is a 1:1 stoichiometric ratio of Ag^+ to I^- ions. If the reaction is carried out with an excess silver nitrate, there will be more Ag^+ than I^- ions in the surface of the particles. The particles will thus be positively charged and the counterions surrounding them will be NO_3^-. If the reaction is carried out using an exact stoichiometric 1:1 ratio of silver nitrate to sodium iodide or with an excess sodium iodide, the surface of the particles will contain an excess I^- over Ag^+ ions.[24,25,27] The particles will be negatively charged, and Na^+ will be the counterions surrounding the particles and neutralizing their charges.

An additional mechanism through which particles acquire electric charges is by the adsorption of ions,[25-27] including ionic surfactants.

Electric Double Layers—The surface layer of a silver iodide particle prepared with an excess of sodium iodide contains more I^- than Ag^+ ions, whereas its bulk contains the two ions in exactly equimolar proportion. The aqueous solution in which this particle is suspended contains relatively high concentrations of Na^+ and NO_3^-, a lower concentration of I^-, and traces of H^+, OH^- and Ag^+.

The negatively charged particle surface attracts positive ions from the solution and repels negative ions: the solution in the vicinity of the surface contains a much higher concentration of Na^+, which are the counterions, and a much lower concentration of NO_3^- ions than the bulk of the solution. A number of Na^+ ions equal to the number of excess I^- ions in the surface (ie, the number of I^- ions in the surface layer minus the number of Ag^+ ions in the surface layer) and equivalent to the net negative surface charge of a particle are pulled towards its surface. These counterions tend to stick to the surface, approaching it as closely as their hydration spheres permit (Helmholtz double layer), but the thermal agitation of the water molecules tends to disperse them throughout the solution. As a result, the layer of counterions surrounding the particle is spread out. The Na^+ concentration is highest in the immediate vicinity of the negative surface, where they form a compact layer called the Stern layer, and decreases with distance from the surface, throughout a diffuse layer called the Gouy-Chapman layer: the sharply defined negatively charged surface is surrounded by a cloud of Na^+ counterions required for electroneutrality. The combination of the two layers of oppositely charged ions constitutes an electric double layer. It is illus-

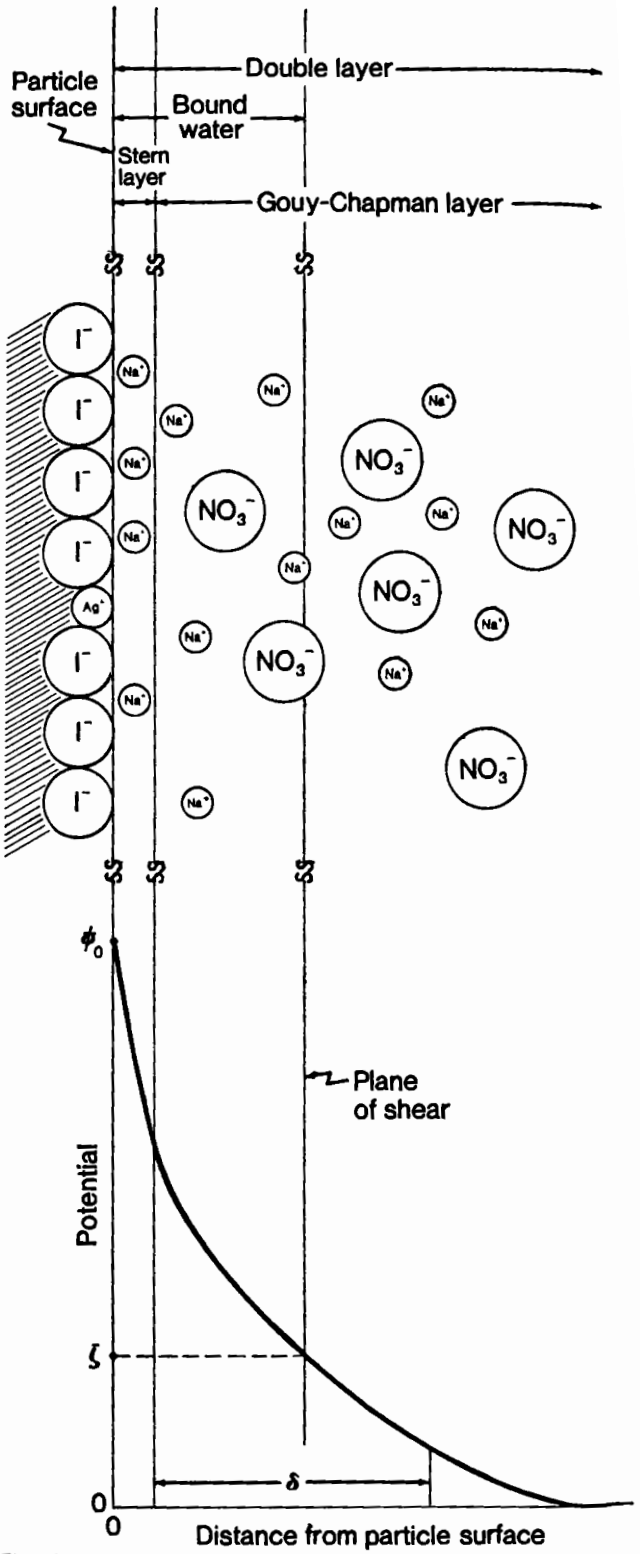


Fig 19-28. Electric double layer at the surface of a silver iodide particle (upper part) and the corresponding potentials (lower part). The distance from the particle surface, plotted on the horizontal axis, refers to both the upper and lower parts.

trated in the top part of Fig 19-28. The horizontal axis represents the distance from the particle surface in both the top and bottom parts.

The electric potential of a plane is equal to the work against electrostatic forces required to bring a unit electric charge from infinity (in this case, from the bulk of the solution) to that plane. If the plane is the surface of the particle, the potential is called surface or ψ_0 potential, which measures the total potential of the double layer. This is the thermodynamic potential which operates in galvanic cells. On moving away from the particle surface towards the bulk solution in the direction of the horizontal axis, the potential drops rapidly across the Stern layer because the Na^+ ions in the immediate vicinity of the surface screen Na^+ ions farther removed, in the diffuse part of the double layer, from the effect of the negative surface charge. The decrease in potential across the Gouy-Chapman layer is more gradual. The diffuse double layer gradually comes to an end as the composition approaches that of the bulk liquid where the anion concentration equals the cation concentration, and the potential approaches zero asymptotically. In view of the indefinite end point, the thickness δ of the diffuse double layer is arbitrarily assigned the value of the distance over which the potential at the boundary between the Stern and Gouy-Chapman layers drops to $1/e = 0.37$ of its value.[24-27] The thickness of double layers usually ranges from 10 to 1000 Å. It decreases as the concentration of electrolytes in solution increases, more rapidly for counterions of higher valence. The value of δ is approximately equal to the reciprocal of the Debye-Hückel theory parameter, κ.

Of practical importance, because it can be measured experimentally, is the electrokinetic or ζ (zeta) potential. In aqueous dispersion, even relatively hydrophobic inorganic particles and organic particles containing polar functional groups are surrounded by a layer of water of hydration attached to them by ion-dipole and dipole-dipole interaction. When a particle moves, this shell of bound water and all ions located inside it move along with the particle. Conversely, if water or a solution flows through a fixed bed of these solid particles, the hydration layer surrounding each particle remains stationary and attached to it. The electric potential at the plane of shear or slip separating the bound water from the free water is the ζ potential. It does not include the Stern layer and only that part of the Gouy-Chapman layer which lies outside the hydration shell. The various potentials are shown on the bottom part of Fig 19-28.

Stabilization by Electrostatic Repulsion—When two uncharged hydrophobic particles are in close proximity, they attract each other by van der Waals secondary valences, mainly by London dispersion forces. For individual atoms and molecules, these forces decrease with the seventh power of the distance between them. In the case of two particles, every atom of one attracts every atom of the other particle. Because the attractive forces are nearly additive, they decay much less rapidly with the interparticle distance as a result of this summation, approximately with the second or third power. Since energies of attraction are equal to force x distance, they decrease approximately with the first or second power of the distance. Therefore, whenever two particles approach each other closely, the attractive forces take over and cause them to adhere. Coagulation occurs as the primary particles aggregate into increasingly larger secondary particles or flocs.

If the dispersion consists of two kinds of particles with positive and negative charges, respectively, the electrostatic attraction between oppositely charged particles is superimposed on the attraction by van der Waals forces, and coagulation is accelerated. If the dispersion contains only one kind, as is customary, all particles have surface charges of the same sign and density. In that case, electrostatic repulsion tends to prevent the particles from approaching closely enough to come within effective range of each other's van der Waals attractive forces, thus stabilizing the dispersion against interparticle attachments or coagulation. The electrostatic repulsive energy has a range of the order of δ.

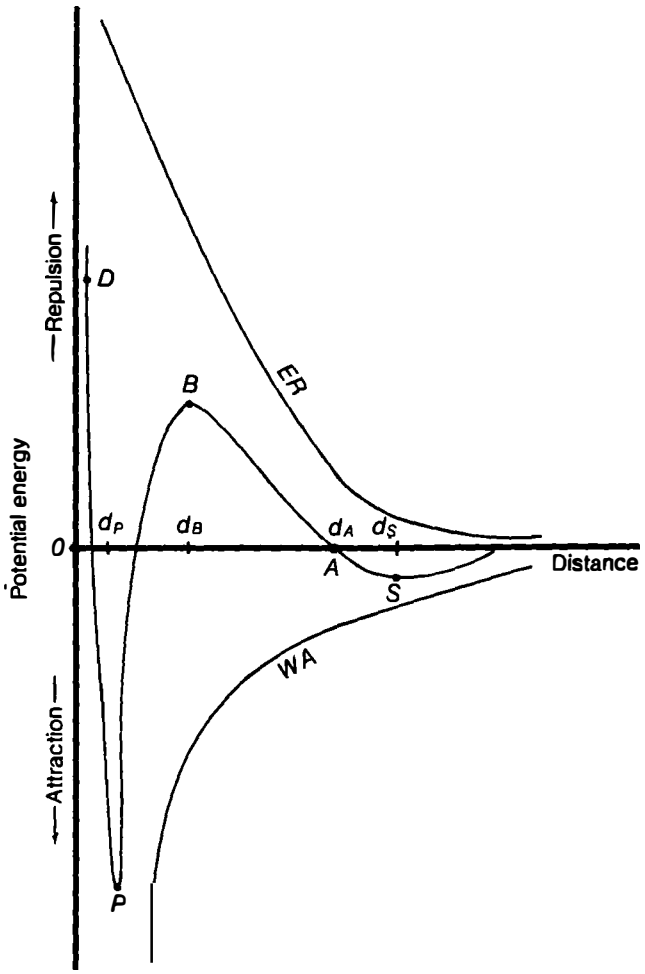


Fig 19-29. Curves representing the van der Waals energy of attraction (*WA*), the energy of electrostatic repulsion (*ER*), and the net energy of interaction (*DPBAS*) between two identical charged particles, as a function of the interparticle distance.

A quantitative theory of the interaction between lyophobic disperse particles was worked out independently by Derjaguin and Landau in the USSR and by Verwey and Overbeek in the Netherlands in the early 1940s.[21,24-27,38] Detailed calculations are also found in Chapter 21 of RPS-17. The so-called DLVO theory predicts and explains many but not all experimental data. Its refinement to account for discrepancies is still continuing.

The DLVO theory is summarized in Fig 19-29, where curve *WA* represents the van der Waals attractive energy which decreases approximately with the second power of the interparticle distance, and curve *ER* represents the electrostatic repulsive energy which decreases exponentially with distance. Because of the combination of these two opposing effects, attraction predominates at small and large distances whereas repulsion may predominate at intermediate distances. Negative energy values indicate attraction, and positive values repulsion. The resultant curve *DPBA*, obtained by algebraic addition of curves *WA* and *ER*, gives the total, net energy of interaction between two particles.

The interparticle attraction depends mainly on the chemical nature and particle size of the material to be dispersed. Once these have been selected, the attractive energy is fixed

and cannot readily be altered. The electrostatic repulsion depends on ψ_0 or the density of the surface charge and on the thickness of the double layer, both of which govern the magnitude of the ζ potential. Thus, stability correlates to some extent with this potential.[24] The ζ potential can be adjusted within wide limits by additives, especially ionic surfactants, water-miscible solvents, and electrolytes (see below). If the absolute value of the ζ potential is small, the resultant potential energy is negative and van der Waals attraction predominates over electrostatic repulsion at all distances. Such sols coagulate rapidly.

The two identical particles whose interaction is depicted in Fig 19-29 have a large (positive or negative) ζ potential resulting in an appreciable positive or repulsive potential energy at intermediate distances. They are on a collision course because of Brownian motion, convection currents, sedimentation, or because the dispersion is being stirred.

As the two particles approach each other, the two atmospheres of counterions surrounding them begin to interpenetrate or overlap at point A corresponding to the distance d_A. This produces a net repulsive (positive) energy because of the work involved in distorting the diffuse double layers and in pushing water molecules and counterions aside, which increases if the particles approach further. If the particles continue to approach each other, even after most of the intervening solution of the counterions between them has been displaced, the repulsion between their surface charges increases the net potential energy of interaction to its maximum positive value at B. If the height of the potential energy barrier B exceeds the kinetic energy of the approaching particles, they will not come any closer than the distance d_B but move away from each other. A net positive potential energy of about 25 kT units usually suffices to keep them apart, rendering the dispersion permanently stable; k is the Boltzmann constant and T is the absolute temperature. At $T = 298°K$, this corresponds to 1×10^{-12} erg. The kinetic energy of a particle is of the order of kT.

On the other hand, if their kinetic energy exceeds the potential energy barrier B, the particles continue to approach each other past d_B, where the van der Waals attraction becomes increasingly more important compared to the electrostatic repulsion. Therefore, the net potential energy of interaction decreases to zero and then becomes negative, pulling the particles still closer together. When the particles touch, at a distance d_P, the net energy has acquired the large negative value P. This deep minimum in potential energy corresponds to a very stable situation in which the particles adhere. Since it is unlikely that enough kinetic energy can be supplied to the particles or that their ζ potential can be increased sufficiently to cause them to climb out of the potential energy well P, they are attached permanently to each other. When most or all of the primary particles agglomerate into secondary particles by such a process, the sol coagulates.

Any closer approach of two particles, than the touching distance d_P, is met with a very rapid rise in potential energy along PD because the solid particles would interpenetrate each other, causing atomic orbitals to overlap (Born repulsion).

Coagulation of Hydrophobic Dispersions—The height of the potential energy barrier and the range over which the electrostatic repulsion is effective (or the thickness of the double layer) determine the stability of hydrophobic dispersions. Both factors are reduced by the addition of electrolytes. The transition between a coagulating and a stable sol is gradual and depends on the time of observation. By using standard conditions, however, it is possible to classify a sol as either coagulated or coagulating, or as stable or fully dispersed.

To determine the value of the coagulating concentration of a given electrolyte for a given sol, a series of test tubes is filled with equal portions of the sol. Identical volumes of solutions of the electrolyte, of increasing concentration, are added with vigorous stirring. After some time at rest (eg, 2 hours), the mixtures are agitated again. After an additional, shorter rest period (eg, ½ hour), they are inspected for signs of coagulation. The tubes can be classified into two groups, one showing no signs of coagulation and the other showing at least some signs, eg, visible flocs. Alternatively, they can be classified into one group showing complete coagulation and the other containing at least some deflocculated colloid left in the supernatant. In either case, the separation between the two classes is quite sharp. The intermediate agitation breaks the weakest interparticle bonds and brings small particles in contact with larger ones, thus increasing the sharpness of separation between coagulation and stability. After repeating the experiment with a narrower range of electrolyte concentrations, the coagulation value c_{CV} of the electrolyte, ie, the lowest concentration at which it coagulates the sol, is established with good reproducibility.[24,25,27]

Typical c_{CV} data for a silver iodide sol prepared with an excess of iodide are listed in Table XIV. The following conclusions can be drawn from the left half of Table XIV:

1. The c_{CV} does not depend on the valence of the anion, since nitrate and sulfate of the same metal have nearly identical values.
2. The differences among the c_{CV}s of cations with the same valence are relatively minor. However, there is a slight but significant trend of decreasing c_{CV} with increasing atomic number in the alkali and in the alkaline earth metal groups. Arranging these cations in the order of decreasing c_{CV} produces the *Hofmeister* or *lyotropic series*. It governs many other colloidal phenomena, including the effect of salts on the temperature of gelation and the swelling of aqueous gels and on the viscosity of hydrosols, the salting out of hydrophilic colloids, the cation exchange on ion-exchange resins, and the permeability of membranes toward salts. The series is also observed in many phenomena involving only small atoms or ions and true solutions, including the ionization potential and electronegativity of metals, the heats of hydration of cations, the size of the hydrated cations, the viscosity, surface tension and infrared spectra of salt solutions, and the solubility of gases therein. For monovalent cations, the lyotropic series is

$$Li^+ > Na^+ > K^+ > NH_4^+ > Rb^+ > Cs^+$$

A similar lyotropic series exists for anions.[21,22,24-26]

The lithium ion has a higher c_{CV} than the cesium ion because it is more extensively hydrated, so that Li^+ (aq), including the hydration shell, is larger than Cs^+ (aq). Owing to its smaller size, the hydrated cesium ion can approach the negative particle surface more closely than the hydrat-

Table XIV—Coagulation Values for Negative Silver Iodide Sol[a]

Electrolyte	c_{CV}, mM/L	Electrolyte	c_{CV}, mM/L
$LiNO_3$	165	$AgNO_3$	0.01
$NaNO_3$	140	½ $(C_{12}H_{25}NH_3)_2SO_4$	0.7
½ Na_2SO_4	141	Strychnine nitrate	1.7
KNO_3	136	½ Morphine sulfate	2.5
½ K_2SO_4	138		
$RbNO_3$	126		
Mean	141		
$Mg(NO_3)_2$	2.60	Quinine sulfate	0.7
$MgSO_4$	2.57		
$Ca(NO_3)_2$	2.40		
$Sr(NO_3)_2$	2.38		
$Ba(NO_3)_2$	2.26		
$Zn(NO_3)_2$	2.50		
$Pb(NO_3)_2$	2.43		
Mean	2.45		
$Al(NO_3)_3$	0.067		
$La(NO_3)_3$	0.069		
$Ce(NO_3)_3$	0.069		
Mean	0.068		

[a] From Ref 21 and unpublished data.

ed lithium ion. Moreover, because of its greater electron cloud, the Cs^+ ion is more polarizable than the Li^+ ion. Therefore, it is more strongly adsorbed in the Stern layer, which makes it a more effective coagulating agent.

3. The coagulation values depend primarily on the valence of the counterions, decreasing by one to two orders of magnitude for each increase of one in their valence (Schulze-Hardy rule). According to the DLVO theory, the coagulation values vary inversely with the sixth power of the valence of the counterions. For mono-, di- and trivalent counterions, they should be in the ratio

$$\frac{1}{1^6} : \frac{1}{2^6} : \frac{1}{3^6} \text{ or } 100 : 1.6 : 0.14$$

The mean c_{CV}'s of Table XIV are 141 : 2.45 : 0.068, or 100 : 1.7 : 0.05, in satisfactory agreement with the DLVO theory.

The following conclusion can be drawn from the right half of Table XIV:

4. The cations on the right side of Table XIV constitute obvious exceptions to the preceding. Ag^+ is the potential-determining counterion. *Potential-determining ions* are those whose concentration determines the surface potential. When silver nitrate is added to the negative silver iodide dispersion, some of its silver ions are incorporated into the negatively charged surface of the particles and lower the magnitude of their charge by reducing the excess of I^- ions in the surface. Thus, silver salts are exceptionally effective coagulating agents because they reduce the magnitude of the ψ_0 as well as of the ζ potential. Indifferent salts, which reduce only the latter, require much higher salt concentrations for comparable reductions in the ζ potential. The other potential-determining ion of silver iodide is I^-. Alkali iodides have higher c_{CV}'s than 141 millimole/liter because they supply iodide ions which enter the surface layer of the silver iodide particles and increase its excess of I^- over Ag^+ ions, thereby making ψ_0 more negative. Bromide and chloride ions act similarly but less effectively.

The principal potential-determining ion for proteins is H^+; those for aluminum hydroxide are OH^- (and hence H^+) and Al^{3+}, but also Fe^{3+} and Cr^{3+} which form mixed hydroxides with Al^{3+}.

5. The cationic surfactant in Table XIV and the alkaloidal salts, which also behave as such, constitute the second exception to the Schulze-Hardy rule. Surface-active compounds contain hydrophilic and hydrophobic moieties in the same molecule, the latter being hydrocarbon portions which by themselves are water-insoluble. Their dual nature causes these compounds to accumulate in interfaces. Dodecylammonium and alkaloidal cations displace inorganic monovalent cations from the Stern layer of a negatively charged silver iodide particle because they are attracted to it not only by electrostatic forces like sodium ions but also by van der Waals forces between their hydrocarbon moieties (dodecyl chains in the case of the dodecylammonium ions) and the solid. Because they are strongly adsorbed from solution onto the surface and do not tend to dissociate from it, surface-active cations are very effective in reducing the ζ potential of the negative silver iodide particles, ie, they have lower c_{CV} than purely inorganic cations of the same valence.

6. Anionic surfactants like those containing lauryl sulfate ions also have a tendency to be adsorbed at solid-liquid interfaces. However, because of electrostatic repulsion between the negatively charged surface of silver iodide particles whose surface layer contains an excess iodide ions and the surface-active anions, adsorption usually does not occur below the critical micelle concentration (see below). If such adsorption does occur, it increases the density of negative charges in the particle surface, raising the c_{CV} of anionic surfactants above that corresponding to their valence.

Ionic solids with surface layers containing the ionic species in near proper stoichiometric balance, and most water-insoluble organic compounds have relatively low surface charge densities. They adsorb ionic surfactants of like charge from solution even at low concentrations, which increases their surface charge densities and the magnitude of their ζ potentials, stabilizing their aqueous dispersions.

The addition of water-miscible solvents such as alcohol, glycerin, propylene glycol or polyethylene glycols to aqueous dispersions lowers the dielectric constant of the medium. This reduces the thickness of the double layer and, therefore, the range over which electrostatic repulsion is effective, and lowers the size of the potential energy barrier. Addition of solvents to aqueous dispersions tends to coagulate them. At concentrations too low to cause coagulation by themselves, solvents make the dispersions more sensitive to coagulation by added electrolytes, ie, they lower the c_{CV}.

Progressive addition of the salt of a counterion of high valence reduces the ζ potential of colloidal particles gradually to zero. Eventually, the sign of the ζ potential may be inverted and its magnitude may increase again, but in the opposite direction. The ψ_0 and ζ potentials of aqueous sulfamerazine suspensions are negative above their isoelectric points; those of bismuth subnitrate are positive. As discussed on page 297, the addition of Al^{3+} to the former and of PO_4^{3-} to the latter in large enough amounts inverts the sign of their ζ potentials; their ψ_0 potentials remain unchanged. Surface-active ions of opposite charge may also produce such charge inversion.

The superposition of the van der Waals attractive energy with its long-range effectiveness and the electrostatic repulsive energy with its intermediate-range effectiveness frequently produces a shallow minimum (designated S in Fig 19-29) in the resultant energy-distance curve at interparticle distances d_S several times greater than δ. If this minimum in potential energy is small compared to kT, Brownian motion prevents aggregation. For large particles such as those of many pharmaceutical suspensions and for particles which are large in one or two dimensions (rods and plates), the *secondary minimum* may be deep enough to trap them at distances d_S from each other. This requires a depth of several kT units. Such fairly long-range and weak attraction produces loose aggregates or flocs which can be dispersed by agitation or by removal or reduction in the concentration of flocculating electrolytes.[21,25–27,38] This reversible aggregation process involving the secondary minimum is called *flocculation*. By contrast, aggregation in the deep primary minimum P, called *coagulation*, is irreversible.

Stabilization by Adsorbed Surfactants—As discussed above, surfactants tend to accumulate at interfaces because of their amphiphilic nature. This process is an *oriented physical adsorption*. Surfactant molecules arrange themselves at the interface between water and an organic solid or liquid of low polarity in such a way that the hydrocarbon chain is in contact with the surface of the solid particle or sticks inside the oil droplet while the polar headgroup is oriented towards the water phase. This orientation removes the hydrophobic hydrocarbon chain from the bulk of the water, where it is unwelcome because it interferes with the hydrogen bonding among the water molecules, while leaving the polar headgroup in contact with water so that it can be hydrated.

Figure 19-30*A* shows schematically that at low surfactant concentration and low surface coverage, the hydrocarbon chains of the adsorbed surfactant molecules lie flat against the solid surface. At higher surfactant concentrations, the surfactant molecules are adsorbed in the upright position to permit the adsorption of more surfactant per unit surface area. Figure 19-30*B* shows a nearly close-packed monolayer of adsorbed surfactant molecules. The terminal methyl groups of their hydrocarbon tails are in contact with the hydrophobic surface and the hydrocarbon tails are in lateral contact with each other. London dispersion forces promote attraction between both types of adjoining groups. The polar headgroups protrude into the water and are hydrated.

The adsorption of ionic surfactants increases the charge density and the ζ potential of the disperse particles. These two parameters are low for organic substances lacking ionic or strongly polar groups. The increase in electrostatic repulsion among the nonpolar organic particles due to adsorption of surface-active ions stabilizes the dispersion against coagulation. This "charge stabilization" is described by the DLVO theory.

Most water-soluble nonionic surfactants are polyoxyethylated (see above): Each molecule consists of a hydrophobic hydrocarbon chain combined with a hydrophilic polyethylene glycol chain, eg $CH_3(CH_2)_{15}(OCH_2CH_2)_{10}OH$. Hydration of the 10 ether groups and of the terminal hydroxyl

group renders the surfactant molecule water-soluble. It adsorbs at the interface between a hydrophobic solid and water, with the hydrocarbon moiety adhering to the solid surface and the polyethylene glycol moiety protruding into the water, where it is hydrated. The particle surface is thus surrounded by a thin layer of hydrated polyethylene glycol chains. This hydrophilic shell forms a steric barrier which prevents close contact between particles and, hence, coagulation ("steric stabilization"). Nonionic surfactants also reduce the sensitivity of hydrophobic dispersions toward coagulation by salts, ie, they increase the coagulation values.[39]

In a flocculated dispersion, groups of several particles are agglomerated into flocs. Frequently, the particles of a floc are in physical contact. When a surfactant is added to a flocculated sol, the dissolved surfactant molecules become adsorbed at the surface of the particles. Surfactant molecules tend to pry apart flocs by wedging themselves between the particles at their areas of contact. This action opens up for surfactant adsorption additional surface area that was previously blocked by adhesion of another solid surface. The breaking up of flocs or secondary particles is defined above as deflocculation or peptization.

Ophthalmic suspensions should be deflocculated because the large particle size of flocs causes eye irritation. Parenteral suspensions should be deflocculated to prevent flocs from blocking capillary blood vessels and hypodermic syringes, and to reduce tissue irritation. Deflocculated suspensions tend to cake, however, ie, the sediment formed by gravitational settling is compact and may be hard to disperse by shaking. Caking in oral suspensions is prevented by controlled flocculation as discussed below.

Stabilization by Adsorbed Polymers—Water-soluble polymers are adsorbed at the interface between water and a hydrophobic solid if they have some hydrophobic groups that limit their water solubility and render them amphiphilic and, hence, surface-active. Such polymers also tend to accumulate at the air-water interface and lower the surface tension of the aqueous phase. A high concentration of ionic groups in polyelectrolytes tends to eliminate surface activity and the tendency to adsorb at interfaces, because the polymer is excessively water-soluble. An example is *sodium carboxymethylcellulose*. *Polyvinyl alcohol* is very water-soluble due to the high concentration of hydroxyl groups and does not adsorb extensively at interfaces. Polyvinyl alcohol is manufactured by the hydrolysis of polyvinyl acetate, which is water-insoluble. Incomplete hydrolysis of, say, only 85% of the acetyl groups produces a copolymer which is water-soluble but surface-active as well. Other surface-active polymers include methylcellulose, hydroxypropyl cellulose, high-molecular-weight polyethylene glycols (polyethylene oxides), and proteins. The surface activity of proteins is due to the presence of hydrophobic groups in the side chains at concentrations too low to cause insolubility in water. Proteins are denatured upon adsorption at air-water and solid-water interfaces.

The long, chain-like polymer molecules are adsorbed from solution onto solid surfaces in the form of loops projecting into the aqueous phase, as shown in Fig 19-31*A*, rather than lying flat against the solid substrate. Only a small portion of the chain segments of an adsorbed macromolecule is actually in contact with and adheres directly to the surface. Because of its great length, however, there are enough of such areas of contact to anchor the adsorbed macromolecule firmly onto the solid. Figure 19-30 is drawn on a much more expanded scale than Fig 19-31.

The sol particles are surrounded by a layer consisting of the adsorbed polymer chains, the water of hydration associated with them, and water trapped mechanically inside the chain loops. This sheath is an integral part of the particle surface. The layers of adsorbed polymer prevent the particles from approaching each other closely enough for the interparticle attraction by London dispersion forces to produce coagulation. These forces are effective only over very small interparticle distances of less than twice the thickness of the adsorbed polymer layer.

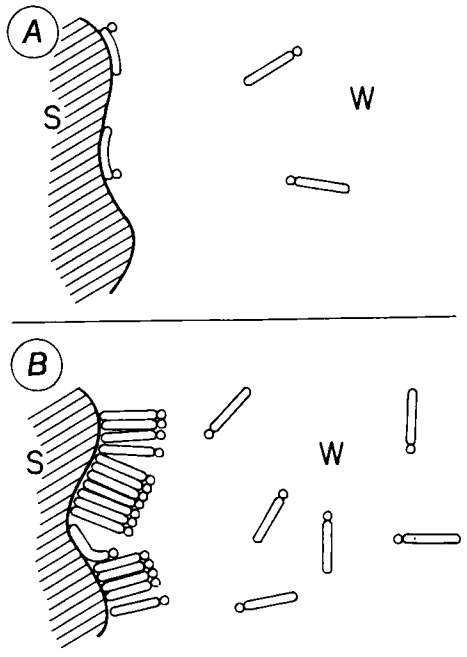


Fig 19-30. Schematic representation of the physical adsorption of surfactant molecules at a hydrophobic solid (S)/water (W) interface. Cylindrical portions and spheres represent hydrocarbon chains and polar headgroups of the surfactant molecules, respectively (*A*) low surfactant concentration/low surface coverage; (*B*) near critical micelle concentration/surface coverage near saturation

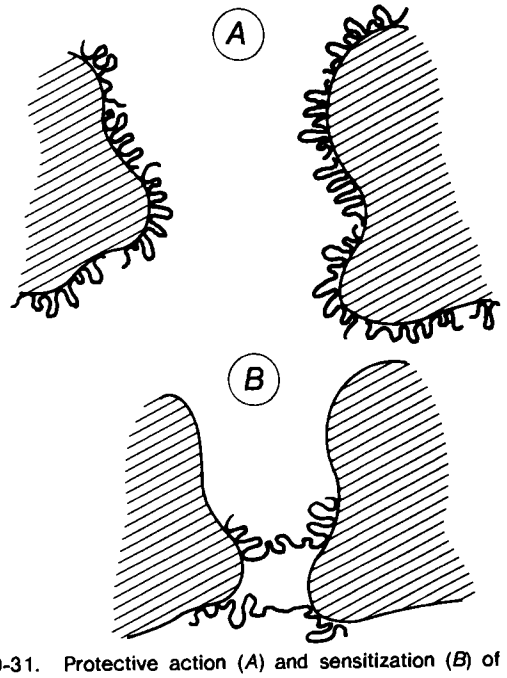


Fig 19-31. Protective action (*A*) and sensitization (*B*) of sols of hydrophobic particles by adsorbed polymer chains.

The mechanisms of *steric stabilization* by which adsorbed nonionic macromolecules prevent coagulation of hydrophobic sols (*protective action*) are also operative in the stabilization of sols by nonionic surfactants. The difference between adsorbed nonionic surfactants and adsorbed polymers

is that the hydrophilic polyethylene glycol moieties of the adsorbed surfactant molecules protruding into water resemble the chain ends of the adsorbed macromolecules rather than their looped segments. The following protective mechanisms are operative:

1. The layer of adsorbed polymer and enmeshed water surrounding the particles forms a *mechanical* or *steric barrier* between them that prevents the close interparticle approach necessary for coagulation. At dense surface coverage, these layers are somewhat elastic. They may be dented by a collision between two particles but tend to spring back.
2. When two particles approach so closely that their adsorbed polymer layers overlap, the chain loops of the two opposing layers compress and mix with or interpenetrate each other. The resulting restriction to the freedom of motion of the chain segments in the overlap region produces a negative entropy change which tends to make the free energy change for the reduction in interparticle distance required for coagulation positive. The reverse process of disentanglement of the two opposing adsorbed polymer layers resulting from separation of the particles occurs because it is energetically more favorable. The particles are thus prevented from coagulation by *entropic repulsion* through the mechanism of *entropic stabilization* of the sol. This mechanism predominates when the concentration of polymer in the adsorbed layer is low.
3. As the polymer layers adsorbed on two approaching particles overlap and compress or interpenetrate each other, more polymer segments become crowded into a given volume of the aqueous region between the particles. The increased polymer concentration in the overlap region causes a local increase in osmotic pressure, which is relieved by an influx of water. This influx to dilute the polymer loops pushes the two particles apart, preventing coagulation.
4. If the adsorbed polymer has some ionic groups, stabilization by electrostatic repulsion or charge stabilization described above is added to the three steric stabilization mechanisms to prevent a close interparticle approach and, hence, coagulation.
5. The adsorption of water-soluble polymers changes the nature of the surface of the hydrophobic particles to hydrophilic, resulting in an increased resistance of the sol to coagulation by salts.[40]

The water-soluble polymers whose adsorption stabilizes hydrophobic sols and protects them against coagulation are called *protective colloids*. *Gelatin* and *serum albumin* are the preferred protective colloids for stabilizing parenteral suspensions because of their biocompatibility. These two polymers, as well as casein (milk protein), dextrin (partially hydrolyzed starch) and vegetable gums like acacia and tragacanth are metabolized in the human body. Cellulose derivatives and most synthetic protective colloids such as *povidone* are not biotransformed. Because of this and because of their large molecular size, polymers pertaining to the last two categories are not absorbed but excreted intact when they are administered in an oral dosage form.

A semiquantitative assessment of the stabilizing efficiency of protective colloids is the *gold number*, developed by Zsigmondy. It is the largest number of milligrams of a protective colloid which, when added to 10 mL of a special standardized gold sol, just fails to prevent the change in color from red to blue on addition of 1 mL of 10% NaCl solution. The gold sol contains 0.0058% gold with a particle size of about 250 Å. Coagulation by sodium chloride causes the color change. Representative gold numbers are 0.005 to 0.01 for gelatin, 0.01 for casein, 0.02 to 0.5 for egg albumin, 0.15 to 0.5 for acacia, and 1 to 7 for dextrin.[22,27] Gelatin is a more effective protective colloid than acacia or dextrin because the presence of some hydrophobic side groups makes it more surface active and causes more extensive adsorption from solution. Other protective numbers are based on different hydrophobic disperse solids, eg, silver, Prussian blue, sulfur, ferric oxide. The ranking of different protective colloids depends somewhat on the substrate. When formulating a disperse dosage form, one should measure the protective action on the actual solid hydrophobic phase to be dispersed as a sol.

Sensitization is the opposite of protective action, namely, a decrease in the stability of hydrophobic sols. It is brought about by some protective colloids, at concentrations well below those at which they exert a protective action. A protective colloid may, at very low concentrations, flocculate a sol in the absence of added salts and/or lower the coagulation values of the sol.

In the case of nonionic polymers or of polyelectrolytes with charges of the same sign as the sol, flocculation is the result of the bridging mechanism illustrated in Fig 19-31*B*. At very low polymer concentrations, there are not nearly enough polymer molecules present to cover each sol particle completely. Since the particle surfaces are largely bare, a single macromolecule may be adsorbed on two particles, bridging the gap between them and pulling them close together. Flocs of several particles are formed when one particle is bridged or connected to two or more other particles by two or more polymer molecules adsorbed jointly on two or possibly even three particles. Such flocculation usually occurs over a narrow range and at very low values of polymer concentrations. At higher concentrations, when enough polymer is available to cover the surface of all particles completely, bridging is unlikely to occur and the adsorbed polymer stabilizes or peptizes the sol.[23,40]

The nonionic Polymer A of Fig 19-32 stabilizes the sol at all concentrations. Neither sensitization by bridging nor by charge neutralization is observed. The reason that Polymer A lowers the positive ζ potential of the sol slightly is that increasing amounts of adsorbed polymer chains gradually shift the plane of shear outward, away from the positively charged surface. If Polymer A was a cationic polyelectrolyte, the ζ potential–protective colloid concentration plot would gradually rise with increasing polymer adsorption rather than drop.

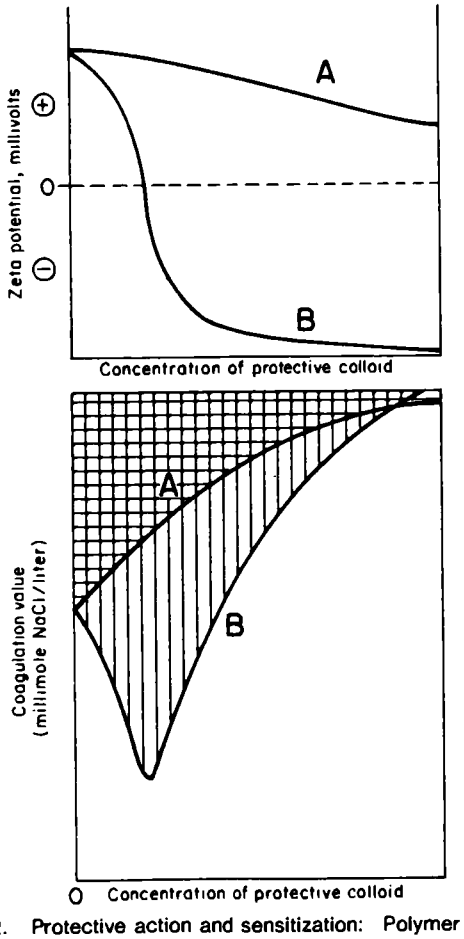


Fig 19-32. Protective action and sensitization: Polymer A exerts protective action at all concentrations, while Polymer B sensitizes at low concentrations and stabilizes at high concentrations. Horizontal and vertical hatching indicates region of flocculation for a sol treated with various concentrations of Polymers A and B, respectively. Clear region underneath indicates sol is deflocculated.

If the polymer has ionic groups of charge opposite to the charge of the sol particles, limited adsorption neutralizes the charge of the particles, reducing their ζ potential to near zero. With stabilization by electrostatic repulsion thus inoperative, and steric stabilization ineffective because of low surface coverage with adsorbed polymer, the sol either coagulates by itself or is coagulated by very small amounts of sodium chloride. At higher polymer concentrations and more extensive adsorption, charge reversal of the particles to the sign of the charge of the polyelectrolyte reactivates charge stabilization and adds steric stabilization, increasing the coagulation value of the sol well above the initial value before polymer addition.

For example, a partly hydrolyzed polyacrylamide with about 20% of ammonium acrylate repeating units is an anionic polyelectrolyte. At the ppm level, the polymer flocculates aluminum hydroxide sols at a pH of 6 to 7, where the sols are positively charged and the polyelectrolyte is fully ionized. At a polymer concentration of 1:10,000, the sol becomes negatively charged because extensive polymer adsorption introduces an excess of $-COO^-$ groups over $=Al^+$ ions into the particle surface. Steric stabilization plus electrostatic repulsion make the sol more stable against flocculation by salts than it was before the polyacrylamide addition.

Polymer B in Fig 19-32 illustrates this example. The curve in the lower plot indicates sensitization, with the coagulation value of sodium chloride lowered by as much as 60%. Zeta potential measurements can distinguish between sensitization by bridging and by charge neutralization. The charge reversal caused by adsorption of Polymer B shown in the upper plot pinpoints charge neutralization as the cause of sensitization. If Polymer B had a ζ potential–polymer concentration plot similar to Polymer A, sensitization would be ascribed to bridging.

Even water-soluble polymers which are too thoroughly hydrophilic to be adsorbed by hydrophobic sol particles can stabilize those sols. Their thickening action slows down Brownian motion and sedimentation, giving the particles less opportunity to come into contact and hence retarding flocculation.

Electrokinetic Phenomena—When a dc electric field is applied to a dispersion, the particles move towards the electrode of charge opposite to that of their surface. The counterions located inside their hydration shell are dragged along while the counterions in the diffuse double layer outside the plane of slip, in the free or mobile solvent, move toward the other electrode. This phenomenon is called *electrophoresis*. If the charged surface is immobile, as is the case with a packed bed of particles or a tube filled with water, application of an electric field causes the counterions in the free water to move towards the opposite electrode, dragging solvent with them. This flow of liquid is called *electroosmosis*, and the pressure produced by it, *electroosmotic pressure*. Conversely, if the liquid is made to flow past charged surfaces by applying hydrostatic pressure, the displacement of the counterions in the free water produces a potential difference between the two ends of the tube or bed called *streaming potential*.

The three phenomena depend on the relative motion of a charged surface and of the diffuse double layer outside the plane of slip surrounding that surface. The major part of the diffuse double layer is within the free solvent and can, therefore, move along the surface.[24—27,41] All three electrokinetic phenomena measure the identical ζ potential, which is the potential at the plane of slip.

The particles of pharmaceutical suspensions and emulsions are visible in the microscope or ultramicroscope, as are bacteria, erythrocytes and other isolated cells, latex particles, and many contaminant particles in pharmaceutical solutions. Their ζ potential is conveniently measured by *microelectrophoresis*. A potential difference E applied between two electrodes dipping into the dispersion and separated by a distance d produces the potential gradient or field strength E/d, expressed in v/cm. From the average velocity v of the particles, measured with the eyepiece micrometer of a microscope and a stopwatch, the ζ potential is calculated by the Smoluchowski equation

$$\zeta = \left(\frac{4\pi\eta}{D}\right)\left(\frac{v}{E/d}\right) = \left(\frac{4\pi\eta}{D}\right)\mu$$

The electrophoretic mobility $u = v/(E/d)$ is the velocity in a potential gradient of 1 v/cm. Particle size and shape do not affect the ζ potential according to the above equation. However, if the particle radius is comparable to δ or smaller (in which case the particles cannot be detected in a microscope), the factor 4 is replaced by 6. The viscosity η and the dielectric constant D refer to the aqueous medium in the double layer and cannot be measured directly.[42] Using the values for water at 25°, expressing the velocity in μm/sec and the electrophoretic mobility in (μm/sec)/(volts/cm), and converting into the appropriate units reduces the Smoluchowski equation to $\zeta = 12.9\,u$, with ζ given in millivolts (mV). If the particle surface has appreciable conductance, the ζ potential calculated by this equation may be low.[25,41,42] Dispersions of hydrophobic particles with ζ potentials below 20–30 mV are frequently unstable and tend to coagulate. On the other hand, values as high as ±180 mV have been reported for the ζ potential.[21,24,41]

The chief experimental precautions in microelectrophoresis measurements are:

1. Electroosmosis causes liquid to flow along the walls of the cell containing the dispersion. This in turn produces a return flow in the center of the cell. The microscope must be focused on the stationary boundary between the two liquid layers flowing in opposite directions in order to measure the true velocity of the particles.
2. Only in very dilute dispersions is it possible to follow the motion of single particles in the microscope field and to measure their velocity. Since the ζ potential depends largely on the nature, ionic strength, and pH of the suspending medium, dispersions should be diluted not with water but with solutions of composition identical to their continuous phase, eg, with their own serum separated by ultrafiltration or centrifugation. The Zeta-Meter is a commercial microelectrophoresis apparatus of easy, fast and reproducible operation.

When the particles cannot be observed individually with a microscope or ultramicroscope, other electrophoresis methods are employed.[24,27,41,43,44] In *moving boundary electrophoresis*, the movement of the boundary formed between a sol or solution and the pure dispersion medium in an electric field is studied. If the disperse phase is colorless, the boundary is located by the refractive index gradient (Tiselius apparatus, used frequently with protein solutions). If several species of particles or solutes with different mobilities are present, each will form a boundary moving with a characteristic velocity. Unlike microelectrophoresis, this method permits the identification of different colloidal components in a mixture, the measurement of the electrophoretic mobility of each, and an estimation of the relative amounts present.

Zone electrophoresis theoretically permits the complete separation of all electrophoretically different components, requires much smaller samples than moving boundary electrophoresis, and can be performed in simpler and less expensive equipment. The method avoids convection by supporting the solution in an inert and porous solid like filter paper, cellulose acetate membrane, agar, starch or polyacrylamide gels cut into strips, or disks or columns of polyacrylamide gel.

A strip of filter paper or gel is saturated with a conducting buffer solution and a few microliters of the solution being analyzed is deposited as a spot or narrow band. A potential difference is applied between the ends of the strip which are

in contact with the electrode compartments. The spot or band spreads and unfolds as each component migrates towards one or the other electrode at a rate determined primarily by its electrophoretic mobility. Evaporation of water due to the heating effect of the electric current may be minimized by immersing the strip in a cooling liquid or sandwiching it between impervious solid sheets. After a sufficient time has elapsed to afford good separation, the strip is removed and dried. The position of the spots or bands corresponding to the individual components is detected by color reactions or radioactive counting.

Zone electrophoresis is applied mainly in analysis and for small-scale preparative separations. It does not permit mobility measurements. Because several samples can be analyzed simultaneously (in parallel strips or gel columns), because only minute amounts of sample are needed, and because the equipment is simple and easy to operate, zone electrophoresis is widely used to study the proteins in blood serum, erythrocytes, lymph and cerebrospinal fluid, saliva, gastric and pancreatic juices and bile.

Immunodiffusion combined with electrophoresis is called *immunoelectrophoresis*.[43,45] The proteins in a fluid, including the antigens, are first separated by gel electrophoresis. A longitudinal trench is then cut along one or both sides of the gel strip near the edge in the direction of the electrophoresis axis. The trench is filled with the antibody solution. On standing, antibody and antigen proteins diffuse in all directions, including toward each other. Precipitation occurs along an elliptical arc (precipitin band) wherever an antigen meets its specific antibody. The precipitin bands are either visible directly or may be developed by staining. Since diseases frequently produce abnormal electrophoretic patterns in body fluids, zone electrophoresis and immunoelectrophoresis are convenient and powerful diagnostic techniques.

Isoelectric focusing[44,46] uses electrophoresis to separate proteins according to their isoelectric points. At pH values equal to their isoelectric points, proteins do not migrate in an electric field because their net charge is zero. In a liquid column on which a pH gradient is imposed, different species arrange themselves so that the protein with the highest isoelectric point will be located nearest to the cathode, which is immersed in the solution of a strong base. The protein with the lowest isoelectric point will be located nearest to the anode, which is immersed in the solution of a strong acid. The other proteins settle into intermediate positions, where the pH values are intermediate and equal to their isoelectric points.

Hydrophilic Dispersions

Most liquid disperse systems of pharmaceutical interest are aqueous. Therefore, most lyophilic colloidal systems discussed below consist of hydrophilic solids dissolved or dispersed in water. Most of the products mentioned below are official in the USP or NF, where more detailed descriptions may be found, also elsewhere in this text.

Hydrophilic colloids can be divided into particulate and soluble materials. The latter are water-soluble linear or branched polymers dissolved molecularly in water. Their aqueous solutions are classified as colloidal dispersions because the individual molecules are in the colloidal particle size range, exceeding 50 or 100 Å. Particulate or corpuscular hydrophilic colloidal dispersions are formed by solids which swell and are peptized in water but whose primary particles do not dissolve or break down into individual molecules or ions. One subdivision of particulate hydrophilic colloids is comprised of dispersions of cross-linked polymers whose linear, uncross-linked analogues are water-soluble.

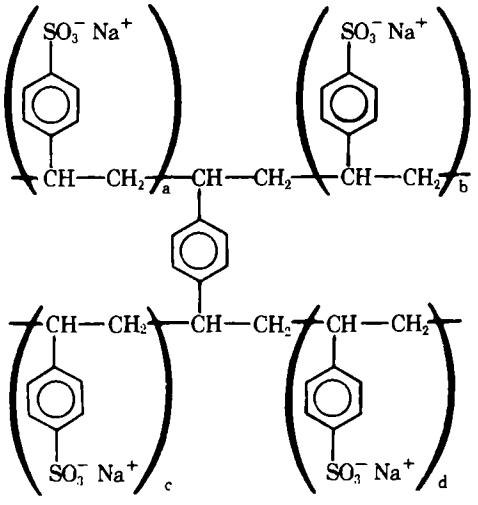
Particulate Hydrophilic Dispersions

The disperse phase of these sols consists of solids which in water swell and break up spontaneously into particles of colloidal dimensions. The disperse particles have high specific surface areas and are, therefore, extensively hydrated. They have characteristic shapes. If the attraction between individual particles is strong, the dispersions have yield values at relatively low solids content.

Bentonite is an aluminum silicate crystallizing in a layer structure (see above), with individual lamellas 9.4 Å thick. Their top and bottom surfaces are sheets of oxygen ions from silica plus an occasional sodium ion neutralizing a silicate ion-exchange site. The clay particles consist of stacks of these lamellas. Water penetrates inside the stacks between lamellas to hydrate the oxygen ions, causing extensive swelling. Bentonite particles in bentonite magma consist of single lamellas and packets of a few lamellas with intercalated water. The specific surface area amounts to several hundred square meters per gram. *Kaolin* also has a layer structure, but does not swell in water because water does not intercalate between individual lattice layers. Kaolin plates dispersed in water are, therefore, much thicker than those of bentonite, *ca* 0.04 to 0.2 μm. In kaolin, hydrated alumina lattice planes alternate with silica planes. Thus, one of the two external surfaces of a kaolin plate consists of a sheet of oxygen ions from silica, the other is a sheet of hydroxide ions from hydrated alumina. Both surfaces are well hydrated. Magnesium aluminum silicate (*Veegum*) is a clay similar to bentonite but contains magnesium; it is white whereas bentonite is gray.

Additional hydrophilic particles producing colloidal dispersions in water are listed below. *Colloidal silicon dioxide* consists of roughly spherical particles covered with siloxane and silanol groups (pages 280–281). *Titanium dioxide* is a white pigment with excellent covering power due to its high refractive index. *Microcrystalline cellulose* (page 279) is hydrophilic because of the hydroxyl and ether groups in the surface of the cellulose crystals. Gelatinous precipitates of hydrophilic compounds such as *aluminum hydroxide gel*, *aluminum phosphate gel*, and *magnesium hydroxide* consist of coarse flocs produced by agglomeration of the colloidal particles formed in the initial stage of the precipitation. They possess large internal surface areas, which is one of the reasons why the first two are used as substrates for adsorbed vaccines and toxoids.

Cross-linked Polymers—The polymers discussed below are polyelectrolytes, ie, they contain ionic groups and would be soluble in water in the absence of cross-linking. For instance, *sodium polystyrene sulfonate* is a copolymer of about 92% styrene and 8% divinylbenzene, which is sulfonated and neutralized to produce the cation-exchange resin



Chains a–b and c–d are water-soluble linear polymer chains. They are cross-linked or bound together via a phenylene group as shown. There are many such cross-links tieing every chain to two or more other chains, so that every atom in a grain of ion-exchange resin is bound to every other atom by primary, covalent bonds. The grains swell in water until the cross-links are strained but do not dissolve, because this would involve the rupture of primary valence bonds. Swelling renders the ion-exchange sites in the interior of a grain accessible to the gastrointestinal fluids. Partial exchange of Na^+ by K^+ followed by excretion of the used resin in the feces reduces hyperkalemia resulting from acute renal failure. Partial replacement of Na^+ by H^+ could reduce acidosis.

Cholestyramine resin is an anion-exchange resin containing the same backbone of cross-linked polystyrene, but substituted with $—CH_2—N^+(CH_3)_3Cl^-$ instead of sodium sulfonate. Part of the chloride anions is exchanged or replaced by bile salt anions, which are thus eliminated in the feces bound to the resin grains rather than reabsorbed. *Colestipol hydrochloride* is another orally administered anion-exchange resin used to increase the fecal excretion of bile salts. It is an extensively cross-linked, insoluble but permeable copolymer made from diethylenetriamine, tetraethylenepentamine, and epichlorohydrin. Strong cation- and anion-exchange resins are used as sustained-release vehicles for basic and acid drugs, respectively (see Chapter 91).

Polycarbophil is a copolymer of acrylic acid cross-linked with a small amount of divinyl glycol. The weakly acidic carboxyl groups are not ionized in the strongly acid environment of the stomach but only in the more nearly neutral intestines. Therefore, swelling by osmotic influx of water occurs mostly in the intestines, where imbibition of water decreases the fluidity of stools associated with diarrhea. Among natural polymers, tragacanth consists of ⅓ of a water-soluble fraction, tragacanthin, and ⅔ of a gel fraction called bassorin which swells in water but does not dissolve. Starch consists of ⅙ of a fraction, soluble in hot water, called amylose. The remainder, amylopectin, merely absorbs water and swells. It owes its insolubility to extensive branching rather than cross-linking.

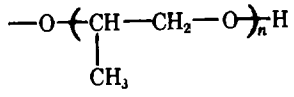
Soluble Polymers as Lyophilic Colloids

Most hydrophilic colloidal systems used in dosage forms are molecular solutions of water soluble, high molecular weight polymers. The polymers are either linear or slightly branched but not cross-linked.

Classifications—According to their origin, water-soluble polymers are divided into three classes. *Natural polymers* include polysaccharides (acacia, agar, heparin sodium, pectin, sodium alginate, tragacanth, xanthan gum) and polypeptides (casein, gelatin, protamine sulfate). Of these, agar and gelatin are only soluble in hot water.

Cellulose derivatives are produced by chemical modification of cellulose obtained from wood pulp or cotton to produce soluble polymers. *Cellulose* is an insoluble, linear polymer of glucose repeat units in the ring or pyranose form joined by β-1,4 glucosidic linkages. Each glucose repeat unit (except for the two terminal ones) contains a primary hydroxyl group on the No 6 carbon and two secondary hydroxyls on No 2 and 3 carbons. The primary hydroxyl is more reactive. Chemical modification of cellulose consists in reactions or substitutions of the hydroxyl groups. The extent of such reactions is expressed as *degree of substitution* (DS), namely, the number of substituted hydroxyl groups per glucose residue. The highest value is DS = 3.0. Fractional values are the rule because the DS is averaged over a multitude of glucose residues. A DS value of 0.6 indicates that some glucose repeat units are unsubstituted while others have one or even two substituents.

Soluble cellulose derivatives are listed below. The DS values correspond to the pharmaceutical grades. The groups shown are the replacements for the hydrogen atoms of the cellulosic hydroxyls. Official derivatives are *methylcellulose* (DS = 1.65–1.93), $—O—CH_3$ and *sodium carboxymethylcellulose* (DS = 0.60–1.00), $—O—CH_2—COO^-Na^+$. *Hydroxyethyl cellulose* (DS ≅ 1.0), $—O\text{-}(CH_2CH_2—O\text{-})_nH$ and *hydroxypropyl cellulose* (DS ≅ 2.5) are manufactured



by the addition of ethylene oxide and propylene oxide, respectively, to alkali-treated cellulose. The value of n is about 2.0 for the former and not much greater than 1.0 for the latter. *Hydroxypropyl methylcellulose* is prepared by reacting alkali-treated cellulose first with methyl chloride to introduce methoxy groups (DS = 1.1–1.8) and then with propylene oxide to introduce propylene glycol ether groups (DS = 0.1–0.3). In general, the introduction of hydroxypropyl groups into cellulose reduces the water solubility somewhat while promoting the solubility in polar organic solvents like short-chain alcohols, glycols and some ethers.

The molecular weight of native cellulose is so high that soluble derivatives of approximately the same degree of polymerization would dissolve too slowly, and their solutions would be excessively viscous even at concentrations of 1% and less. Controlled degradation is used to break the cellulose chains into shorter segments, reducing the viscosity of the solutions of the corresponding soluble derivatives. Commercial grades of a given cellulose derivative such as sodium carboxymethylcellulose come in various molecular weights or viscosity grades as well as with various degrees of substitution, offering the pharmacist a wide selection.

Official cellulose derivatives which are insoluble in water but soluble in some organic solvents include *ethylcellulose* (DS = 2.2–2.7), $—O—C_2H_5$; *cellulose acetate phthalate* (DS = 1.70 for acetyl and 0.77 for phthalyl); and *pyroxylin* or cellulose nitrate (DS ≅ 2), $—O—NO_2$. *Collodion*, a 4.0% w/v solution of pyroxylin in a mixture of 75% (v/v) ether and 25% (v/v) ethyl alcohol, constitutes a lyophilic colloidal system.

The third class, water soluble *synthetic polymers*, consists mostly of vinyl derivatives including *polyvinyl alcohol, povidone* or polyvinylpyrrolidone, and *carbomer* (*Carbopol*), a copolymer of acrylic acid. High molecular weight polyethylene glycols are also called *polyethylene oxides*.

A second classification of hydrophilic polymers is based on their charge. *Nonionic* or uncharged polymers include methylcellulose, hydroxyethyl and hydroxypropyl cellulose, ethylcellulose, pyroxylin, polyethylene oxide, polyvinyl alcohol and povidone. *Anionic* or negatively charged *polyelectrolytes* include the following carboxylated polymers: acacia, alginic acid, pectin, tragacanth, xanthan gum and carbomer at pH values leading to ionization of the carboxyl groups; sodium alginate and sodium carboxymethylcellulose; also polypeptides at pH values above their isoelectric points, eg, sodium caseinate. A stronger acid group is sulfuric acid, which exists as a monoester in agar and heparin and as a monoamide in heparin. *Cationic* or positively charged *polyelectrolytes* are rare. Examples are polypeptides at pH values below their isoelectric points. Protamines are strongly basic due to a high arginine content, with isoelectric points around pH 12, eg protamine sulfate.

Gel Formation—As described in Chapter 20 and illustrated in Fig 20-7*A*, the flexible chains of dissolved polymers interpenetrate and are entangled because of the constant Brownian motion of their segments. The chains writhe and forever change their conformations. Each chain is encased in a sheath of solvent molecules that solvate its functional groups. In the case of aqueous solutions, water molecules

are hydrogen-bonded to the hydroxyl groups of polyvinyl alcohol, hydroxyl groups and ether links of polysaccharides, ether links of polyethylene oxide or polyethylene glycol, amide groups of polypeptides and povidone, and carboxylate groups of anionic polyelectrolytes. The envelope of water of hydration prevents chains segments in close proximity from touching and attracting one another by interchain hydrogen bonds and van der Waals forces as they do in the solid state. The slippage of solvated chains past one another when the solution flows is lubricated by the free solvent between their solvation sheaths.

Factors that lower the hydration of dissolved macromolecules reduce or thin out the sheath of hydration separating adjacent chains. When the hydration is low, contiguous chains tend to attract one another by secondary valence forces including hydrogen bonds and van der Waals forces. Hydrophobic bonding makes an important contribution to interchain attraction between polypeptide chains even in solution. Van der Waals forces and hydrogen bonds thus establish weak and reversible cross-links between chains at their points of contact or entanglement, bringing about phase separation or precipitation.

Most water-soluble polymers have higher solubilities in hot than in cold water and tend to precipitate on cooling, as the sheaths of hydration surrounding adjacent chains become too sparse to prevent interchain attraction. Dilute solutions separate into a solvent phase practically free of polymer and a viscous liquid phase containing practically all of the polymer but still a large excess of solvent. This process is called *simple coacervation* and the polymer-rich liquid phase a *coacervate*.[21,47] If the polymer solution is concentrated enough and/or the temperature low enough, cooling causes the formation of a continuous network of precipitating chains attached to one another through weak cross-links consisting of interchain hydrogen bonds and van der Waals forces at the points of mutual contact. Segments of regularly sequenced polymer chains even associate laterally into crystalline bundles or crystallites. Irregular chain structures as found in random copolymers, randomly substituted cellulose ethers and esters, and highly branched polymers like acacia prevent crystallization during precipitation from solution. Chain entanglements provide the sole temporary cross-links in those cases. The network of associated polymer chains immobilizes the solvent and causes the solution to set to a gel. Gelatinous precipitates or highly swollen flocs may separate when cooling more dilute polymer solutions.

Besides the chemical nature of polymer and solvent, the three most important factors causing phase separation, precipitation and gelation of polymer solutions are temperature, concentration and molecular weight. Lower temperatures, higher concentrations and higher molecular weights promote gelation and produce stronger gels.

For a typical *gelatin*, 10% solutions acquire yield values and begin to gel at about 25°, 20% solutions at about 30° and 30% solutions at about 32°. The *gelation* is reversible: the gels liquefy when heated above these temperatures. Gelation is rarely observed above 34° regardless of concentration, so that gelatin solutions do not gel at 37°. Conversely, gelatin will dissolve readily in water at body temperature. The gelation temperature or gel point of gelatin is highest at the isoelectric point, where the attachment between adjacent chains by coulombic attraction or ionic bonds between carboxylate ions and alkylammonium, guanidinium or imidazolium groups is most extensive. Since the carboxyl groups are not ionized at gastric pH, interchain ionic bonds are practically nonexistent, and interchain attraction is limited to hydrogen bonds and van der Waals forces. The gelation temperature or the melting point of gelatin gels depends more strongly on temperature and concentration than on pH.[48,49] The combination of an acid pH considerably below the isoelectric point and a temperature of 37° completely prevents the gelation of gelatin solutions. Conversely, these two conditions promote rapid dissolution of gelatin capsules in the stomach. Agar and pectic acid solutions set to gels at only a few percent of solids.

Unlike most water-soluble polymers, methylcellulose, hydroxypropyl cellulose and polyethylene oxide are more soluble in cold than in hot water. Their solutions therefore tend to gel on heating (*thermal gelation*).

When dissolving powdered polymers in water, temporary gel formation often slows the process down considerably. As water diffuses into loose clumps of powder, their exterior frequently turns to a cohesive gel of solvated particles encasing dry powder. Such blobs of gel dissolve very slowly because of their high viscosity and the low diffusion coefficient of the macromolecules. Especially for large-scale dissolution, it is helpful to disperse the polymer powder in water before it can agglomerate into lumps of gel. In order to permit dispersion to precede hydration and to prevent temporary gel formation, the polymer powders are dispersed in water at temperatures where the solubility of the polymer is lowest. Most polymer powders, such as sodium carboxymethylcellulose, are dispersed with high shear in *cold* water before the particles can hydrate and swell to sticky gel grains agglomerating into lumps. Once the powder is well dispersed, the solution is heated with moderate shear to about 60° for fastest dissolution. Because methylcellulose hydrates most slowly in hot water, the powder is dispersed with high shear in 1/5 to 1/3 of the required amount of water heated to 80 to 90°. Once the powder is finely dispersed, the rest of the water is added cold or even as ice, and moderate stirring causes prompt dissolution. For maximum clarity, fullest hydration and highest viscosity, the solution should be cooled to 0 to 10° for about an hour.

The following are two alternative methods for preventing the formation of gelatinous lumps upon addition of water. The powder is prewetted with a water-miscible organic solvent such as ethyl alcohol or propylene glycol that does not swell the polymer, in the proportion of from three to five parts solvent to each part of polymer. If other nonpolymeric powdered adjuvants are to be incorporated into the solution, these are dry-blended with the polymer powder. The latter should comprise 1/4 or less of the blend for best results.

A pharmaceutical application of *gelation* in a nonaqueous medium is the manufacture of *Plastibase* or *Jelene* (*Squibb*), which consists of 5% of a low-molecular-weight polyethylene and 95% of mineral oil. The polymer is soluble in mineral oil above 90°, which is close to its melting point. When the solution is cooled below 90°, the polymer precipitates and causes gelation. The mineral oil is immobilized in the network of entangled, and adhering, insoluble polyethylene chains which probably even associate into small crystalline regions. Unlike petrolatum, this gel can be heated to about 60° without substantial loss in consistency.

Large increases in the concentration of polymer solutions may lead to precipitation and gelation. One way of effectively increasing the concentration of aqueous polymer solutions is to add inorganic salts. The salts will bind part of the water of the polymer solution in order to become hydrated. Competition for water of hydration dehydrates the polymer molecules and precipitates them, causing gelation. This phenomenon is called *salting out*. Because of its high solubility in water, ammonium sulfate is often used by biochemists to precipitate and separate proteins from dilute solution. To the pharmacist, salting out usually represents an undesirable problem. It is reversible, however, and subsequent addition of water redissolves the precipitated polymers and liquefies their gels. Salting out may cause the polymer to separate as a concentrated and viscous liquid solution or simple coacervate rather than as a solid gel.

The effectiveness of electrolytes to salt out, precipitate or gel hydrophilic colloidal systems depends on how extensively the electrolytes are hydrated. The *Hofmeister* or *lyotropic series* arranges ions in the order of increasing hydration and increasing effectiveness in salting out hydrophilic colloids. The series, for monovalent cations, is

$$Cs^+ < Rb^+ < NH_4^+ < K^+ < Na^+ < Li^+$$

and for divalent cations,

$$Ba^{2+} < Sr^{2+} < Ca^{2+} < Mg^{2+}$$

This series also arranges the cations in the order of decreasing coagulating power or increasing coagulation values for negative hydrophobic sols (see Table XIV) and of increasing ease of their displacement from cation exchange resins: K^+ displaces Na^+ and Li^+. For anions, the lyotropic series in the order of decreasing coagulating power and decreasing effectiveness in salting out is

$$F^- > \text{citrate}^{3-} > HPO_4^{2-} > \text{tartrate}^{2-} > SO_4^{2-} > \text{acetate}^- > Cl^- > NO_3^- > ClO_3^- > Br^- > ClO_4^- > I^- > CNS^-$$

Iodides and thiocyanates and to a lesser extent bromides and nitrates actually tend to increase the solubility of polymers in water, salting them in.[21,22,24-26] These large polarizable anions destructure water, reducing the extent of hydrogen bonding among water molecules and thereby making more of the hydrogen-bonding capacity of water available to the solute. Most salts except nitrates, bromides, perchlorates, iodides and thiocyanates raise the temperature of precipitation or gelation of most hydrophilic colloidal solutions or their gel melting points. Exceptions among hydrophilic colloids are methylcellulose, hydroxypropyl cellulose and polyethylene oxide whose gelation temperatures or gel points and gel melting points are lowered by salting out.

Hydrophobic aqueous dispersions are coagulated by electrolytes at 0.0001–0.1 *M* concentrations (see Table XIV). Moreover, the coagulation is irreversible. ie, removal of the coagulating salt does not allow the coagulum to be redispersed, because the hydrophobic sols are intrinsically unstable. By contrast, most hydrophilic sols require electrolyte concentrations of 1 *M* or higher for precipitation. Their precipitation or gelation can be reversed, and the polymer redissolved by removing the salt through dialysis or by adding more water. Hydrophilic colloids disperse or dissolve spontaneously in water, and their sols are intrinsically stable.

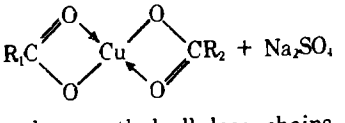
Most of the hydrophilic and water-soluble polymers mentioned above are only slightly soluble or insoluble in alcohol. Addition of alcohol to their aqueous solutions may cause precipitation or gelation because alcohol is a nonsolvent or precipitant, lowering the dielectric constant of the medium, and it tends to dehydrate the hydrophilic solute. Alcohol lowers the concentrations at which electrolytes salt out hydrophilic colloids. Phase separation through the addition of alcohol to an aqueous polymer solution may cause coacervation, ie, the separation of a concentrated viscous liquid phase, rather than precipitation or formation of a gel. Sucrose also competes for water of hydration with hydrophilic colloids, and may cause phase separation. However, most hydrophilic sols tolerate substantially higher concentrations of sucrose than of electrolytes or alcohol. Lower viscosity grades of a given polymer are usually more resistant to electrolytes, alcohol and sucrose than grades of higher viscosity and higher molecular weights.

Whenever hydrophilic colloidal dispersions undergo irreversible precipitation or gelation, chemical reactions are involved. Neither dilution with water nor heating nor attempts to remove the gelling or precipitating agent by washing or dialysis will liquefy those gels or redissolve the gelatinous precipitates formed at lower polymer concentrations. Carboxyl groups are not ionized in strongly acid media. If a polymer owes its solubility to the ionization of these weakly acid groups, reducing the pH of its solution below 3 may lead to precipitation or gelation. This is observed with such carboxylated polymers as many gums, sodium carboxymethylcellulose and carbomer. Hydrogen carboxymethylcellulose swells and disperses but does not dissolve in water. Neutralization to higher pH values returns the carboxyl groups to their ionized state and reverses the gelation or precipitation.

Only the sodium, potassium, ammonium and triethanolammonium salts of carboxylated polymers are well soluble in water. In the case of carboxymethylcellulose, salts with heavy metal cations (silver, copper, mercury, lead) and trivalent cations (aluminum, chromic, ferric) are practically insoluble. Salts with divalent cations, especially of the alkaline earth metals, have borderline solubilities. Generally, higher degrees of substitution tend to increase the tolerance of the carboxymethylcellulose to salts.

Precipitation or gelation occur due to metathesis when inorganic salts of heavy or trivalent cations are mixed with alkali metal salts of carboxylated polymers in solution. For instance, if a soluble copper salt is added to a solution of sodium carboxymethylcellulose, the double decomposition can be written schematically as

$$R_1COO^-Na^+ + R_2COO^-Na^+ + CuSO_4 \longrightarrow$$



$+ Na_2SO_4$

R_1 and R_2 represent two carboxymethylcellulose chains which are cross-linked by a chelated copper ion. Dissociation of the cupric carboxylate complex is negligible.

Particle Phenomena and Coarse Dispersions

The Dispersion Step

The pharmaceutical formulator is concerned primarily with producing a smooth, uniform, easily flowing (pouring or spreading) suspension or emulsion in which dispersion of particles can be effected with minimum expenditure of energy.

In preparing suspensions, particle–particle attractive forces need to be overcome by the high shearing action of such devices as the colloid mill, or by use of surface-active agents. The latter greatly facilitate wetting of lyophobic powders and assist in the removal of surface air that shearing alone may not remove; thus the clumping tendency of the particles is reduced. Moreover, lowering of the surface free energy by the adsorption of these agents directly reduces the thermodynamic driving force opposing dispersion of the particles.

In emulsification shear rates are frequently necessary for dispersion of the internal phase into fine droplets. The shear forces are opposed by forces operating to resist distortion and subsequent breakup of the droplets. Again surface-active agents help greatly by lowering interfacial ten-

sion, which is the primary reversible component resisting droplet distortion. Surface-active agents also may play an important role in determining whether an oil-in-water or a water-in-oil emulsion preferentially survives the shearing action.

Once the process of dispersion begins there develops simultaneously a tendency for the system to revert to an energetically more stable state, manifested by flocculation, coalescence, sedimentation, crystal growth, and caking phenomena. If these physical changes are not inhibited or controlled, successful dispersions will not be achieved or will be lost during shelf life.

Settling and Its Control

In order to control the settling of dispersed material in suspension, the pharmacist must be aware of those physical factors that will affect the rate of sedimentation of particles under ideal and nonideal conditions. He must also be aware of the various coefficients used to express the amount of flocculation in the system and the effect flocculation will have on the structure and volume of the sediment.

Sedimentation Rate

The rate at which particles in a suspension sediment is related to their size and density and the viscosity of the suspension medium. Brownian movement may exert a significant effect, as will the absence or presence of flocculation in the system.

Stokes' Law—The velocity of sedimentation of a uniform collection of spherical particles is governed by Stokes' law, expressed as follows:

$$v = \frac{2r^2(\rho_1 - \rho_2)g}{9\eta} \qquad (35)$$

where v is the terminal velocity in cm/sec, r is the radius of the particles in cm, ρ_1 and ρ_2 are the densities (g/cm^3) of the dispersed phase and the dispersion medium, respectively, g is the acceleration due to gravity (980.7 cm/sec^2) and η is the Newtonian viscosity of the dispersion medium in poises (g/cm sec). Stokes' law holds only if the downward motion of the particles is not sufficiently rapid to cause turbulence. Micelles and small phospholipid vesicles do not settle unless they are subjected to centrifugation.

While conditions in a pharmaceutical suspension are not in strict accord with those laid down for Stokes' law, Eq 35, provides those factors that can be expected to influence the rate of settling. Thus, sedimentation velocity will be reduced by decreasing the particle size, provided the particles are kept in a deflocculated state. The rate of sedimentation will be an inverse function of the viscosity of the dispersion medium. However, too high a viscosity is undesirable, especially if the suspending medium is Newtonian rather than shear-thinning (see Chapter 20), since it then becomes difficult to redisperse material which has settled. It also may be inconvenient to remove a viscous suspension from its container. When the size of particles undergoing sedimentation is reduced to approximately 2 μm, random Brownian movement is observed and the rate of sedimentation departs markedly from the theoretical predictions of Stokes' law. The actual size at which Brownian movement becomes significant depends on the density of the particle as well as the viscosity of the dispersion medium.

Flocculation and Deflocculation—Zeta potential ψ_z is a measurable indication of the potential existing at the surface of a particle. When ψ_z is relatively high (25 mV or more), the repulsive forces between two particles exceed the attractive London forces. Accordingly, the particles are dispersed and are said to be *deflocculated*. Even when brought close together by random motion or agitation, deflocculated particles resist collision due to their high surface potential.

The addition of a preferentially adsorbed ion whose charge is opposite in sign to that on the particle leads to a progressive lowering of ψ_z. At some concentration of the added ion the electrical forces of repulsion are lowered sufficiently that the forces of attraction predominate. Under these conditions the particles may approach each other more closely and form loose aggregates, termed flocs. Such a system is said to be *flocculated*.

Some workers restrict the term *flocculation* to the aggregation brought about by chemical bridging; aggregation involving a reduction of repulsive potential at the double layer is referred to as *coagulation*. Other workers regard flocculation as aggregation in the secondary minimum of the potential energy curve of two interacting particles and coagulation as aggregation in the primary minimum. In the present chapter the term *flocculation* is used for all aggregation processes, irrespective of mechanism.

The continued addition of the flocculating agent can reverse the above process, if the zeta potential increases sufficiently in the opposite direction. Thus, the adsorption of anions onto positively charged deflocculated particles in suspension will lead to flocculation. The addition of more anions can eventually generate a net negative charge on the particles. When this has achieved the required magnitude, deflocculation may occur again. The only difference from the starting system is that the net charge on the particles in their deflocculated state is negative rather than positive.

Table XV—Relative Properties of Flocculated and Deflocculated Particles in Suspension

Deflocculated	Flocculated
1. Particles exist in suspension as separate entities.	Particles form loose aggregates.
2. Rate of sedimentation is slow, since each particle settles separately and particle size is minimal.	Rate of sedimentation is high, since particles settle as a floc, which is a collection of particles.
3. A sediment is formed slowly.	A sediment is formed rapidly.
4. The sediment eventually becomes very closely packed, due to weight of upper layers of sedimenting material. Repulsive forces between particles are overcome and a hard cake is formed which is difficult, if not impossible, to redisperse.	The sediment is loosely packed and possesses a scaffold-like structure. Particles do not bond tightly to each other and a hard, dense cake does not form. The sediment is easy to redisperse, so as to reform the original suspension.
5. The suspension has a pleasing appearance, since the suspended material remains suspended for a relatively long time. The supernatant also remains cloudy, even when settling is apparent.	The suspension is somewhat unsightly, due to rapid sedimentation and the presence of an obvious, clear supernatant region. This can be minimized if the volume of sediment is made large. Ideally, volume of sediment should encompass the volume of the suspension.

Some of the major differences between suspensions of flocculated and deflocculated particles are presented in Table XV.

Effect of Flocculation—In a deflocculated system containing a distribution of particle sizes, the larger particles naturally settle faster than the smaller particles. The very small particles remain suspended for a considerable length of time, with the result that no distinct boundary is formed between the supernatant and the sediment. Even when a sediment becomes discernible, the supernatant remains cloudy.

When the same system is flocculated (in a manner to be discussed later), two effects are immediately apparent. First, the flocs tend to fall together so that a distinct boundary between the sediment and the supernatant is readily observed; second, the supernatant is clear, showing that the very fine particles have been incorporated into the flocs. The initial rate of settling in flocculated systems is determined by the size of the flocs and the porosity of the aggregated mass. Under these circumstances it is perhaps better to use the term *subsidence*, rather than sedimentation.

Quantitative Expressions of Sedimentation and Flocculation

Frequently, the pharmacist needs to assess a formulation in terms of the amount of flocculation in the suspension and to compare this with that found in other formulations. The two parameters commonly used for this purpose are outlined below.

Sedimentation Volume—The *sedimentation volume*, F, is the ratio of the equilibrium volume of the sediment, V_u, to the total volume of the suspension, V_0. Thus,

$$F = V_u/V_0 \tag{36}$$

As the volume of suspension which appears occupied by the sediment increases, the value of F, which normally ranges from nearly 0 to 1, increases. In the system where $F = 0.75$, for example, 75% of the total volume in the container is apparently occupied by the loose, porous flocs forming the sediment. This is illustrated in Fig 19-33. When $F = 1$, no sediment is apparent even though the system is flocculated. This is the ideal suspension for, under these conditions, no sedimentation will occur. Caking also will be absent. Furthermore, the suspension is esthetically pleasing, there being no visible, clear supernatant.

Degree of Flocculation—A better parameter for comparing flocculated systems is the *degree of flocculation*, β, which relates the sedimentation volume of the flocculated suspension, F, to the sedimentation volume of the suspension when deflocculated, F_∞. It is expressed as

$$\beta = F/F_\infty \tag{37}$$

The degree of flocculation is, therefore, an expression of the increased sediment volume resulting from flocculation.

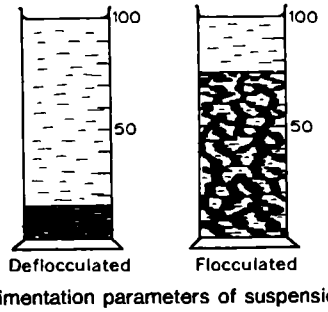


Fig 19-33. Sedimentation parameters of suspensions. Deflocculated suspension: $F_\infty = 0.15$ Flocculated suspension: $F = 0.75$; $\beta = 5.0$.

If, for example, β has a value of 5.0 (Fig 19-33), this means that the volume of sediment in the flocculated system is five times that in the deflocculated state. If a second flocculated formulation results in a value for β of say 6.5, this latter suspension obviously is preferred, if the aim is to produce as flocculated a product as possible. As the degree of flocculation in the system decreases, β approaches unity, the theoretical minimum value.

Suspensions and their Formulation

A pharmaceutical suspension may be defined as a coarse dispersion containing finely divided insoluble material suspended in a liquid medium. Suspension dosage forms are given by the oral route, injected intramusculary or subcutaneously, applied to the skin in topical preparations, and used ophthalmically in the eye. They are an important class of dosage form. Since some products are occasionally prepared in a dry form, to be placed in suspension at the time of dispensing by the addition of an appropriate vehicle, this definition is extended to include these products.

There are certain criteria that a well-formulated suspension should meet. The dispersed particles should be of such a size that they do not settle rapidly in the container. However, in the event that sedimentation occurs, the sediment must not form a hard cake. Rather, it must be capable of redispersion with a minimum effort on the part of the patient. Additionally, the product should be easy to pour, pleasant to take, and resistant to microbial attack.

The three major problem areas associated with suspensions are (1) adequate dispersion of the particles in the vehicle, (2) settling of the dispersed particles, and (3) caking of these particles in the sediment so as to resist redispersion. Much of the following discussion will deal with the factors that influence these processes and the ways in which they can be minimized.

The formulation of a suspension possessing optimal physical stability depends on whether the particles in suspension are to be flocculated or to remain deflocculated. One approach involves use of a structured vehicle to keep deflocculated particles in suspension; a second depends on controlled flocculation as a means of preventing cake formation. A

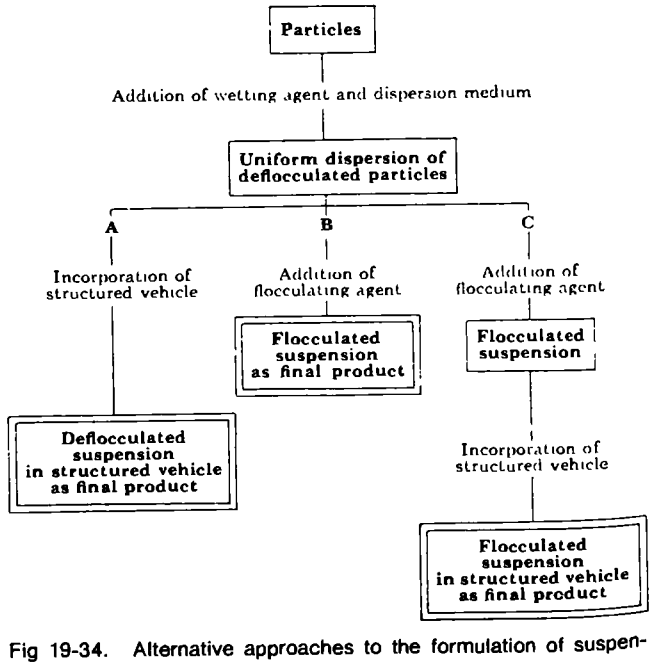


Fig 19-34. Alternative approaches to the formulation of suspensions.

third, a combination of the two previous methods, results in a product with optimum stability. The various schemes are illustrated in Fig 19-34.

Dispersion of Particles—The dispersion step has been discussed earlier in this chapter. Surface-active agents commonly are used as wetting agents; maximum efficiency is obtained when the HLB value lies within the range of 7 to 9. A concentrated solution of the wetting agent in the vehicle may be used to prepare a slurry of the powder; this is diluted with the required amount of vehicle. Alcohol and glycerin may be used sometimes in the initial stages to disperse the particles, thereby allowing the vehicle to penetrate the powder mass.

Only the minimum amount of wetting agent should be used, compatible with producing an adequate dispersion of the particles. Excessive amounts may lead to foaming or impart an undesirable taste or odor to the product. Invariably, as a result of wetting, the dispersed particles in the vehicle are deflocculated.

Structured Vehicles—Structured vehicles are generally aqueous solutions of polymeric materials, such as the hydrocolloids, which are usually negatively charged in aqueous solution. Typical examples are methylcellulose, carboxymethylcellulose, bentonite, and Carbopol. The concentration employed will depend on the consistency desired for the suspension which, in turn, will relate to the size and density of the suspended particles. They function as viscosity-imparting suspending agents and, as such, reduce the rate of sedimentation of dispersed particles.

The rheological properties of suspending agents are considered elsewhere (Chapter 20). Ideally, these form pseudoplastic or plastic systems which undergo shear-thinning. Some degree of thixotropy is also desirable. Non-Newtonian materials of this type are preferred over Newtonian systems because, if the particles eventually settle to the bottom of the container, their redispersion is facilitated by the vehicle thinning when shaken. When the shaking is discontinued, the vehicle regains its original consistency and the redispersed particles are held suspended. This process of redispersion, facilitated by a shear-thinning vehicle, presupposes that the deflocculated particles have not yet formed a cake. If sedimentation and packing have proceeded to the point where considerable caking has occurred, redispersion is virtually impossible.

Controlled Flocculation—When using this approach (see Fig 19-34, B and C), the formulator takes the deflocculated, wetted dispersion of particles and attempts to bring about flocculation by the addition of a flocculating agent; most commonly, these are either electrolytes, polymers, or surfactants. The aim is to *control* flocculation by adding that amount of flocculating agent which results in the maximum sedimentation volume.

Electrolytes are probably the most widely used flocculating agents. They act by reducing the electrical forces of repulsion between particles, thereby allowing the particles to form the loose flocs so characteristic of a flocculated suspension. Since the ability of particles to come together and form a floc depends on their surface charge, zeta potential measurements on the suspension, as an electrolyte is added, provide valuable information as to the extent of flocculation in the system.

This principle is illustrated by reference to the following example, taken from the work of Haines and Martin.[50] Particles of sulfamerazine in water bear a negative charge. The serial addition of a suitable electrolyte, such as aluminum chloride, causes a progressive reduction in the zeta potential of the particles. This is due to the preferential adsorption of the trivalent aluminum cation. Eventually, the zeta potential will reach zero and then become positive as the addition of $AlCl_3$ is continued.

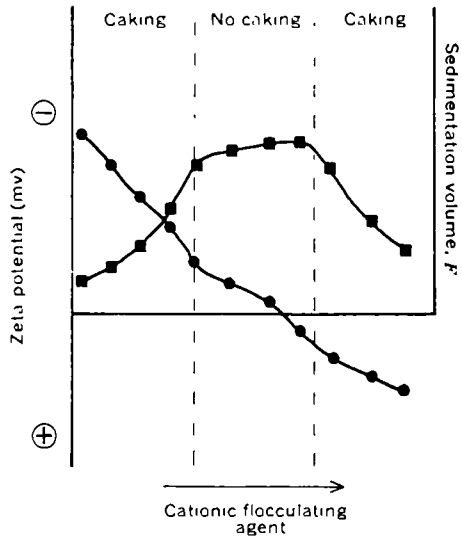


Fig 19-35. Typical relationship between caking, zeta potential and sedimentation volume, as a positively charged flocculating agent is added to a suspension of negatively charged particles. ●: zeta potential; ■: sedimentation volume

If sedimentation studies are run simultaneously on suspensions containing the same range of $AlCl_3$ concentrations, a relationship is observed (Fig 19-35) between the sedimentation volume, F, the presence or absence of caking, and the zeta potential of the particles. In order to obtain a flocculated, noncaking suspension with the maximum sedimentation volume, the zeta potential must be controlled so as to lie within a certain range (generally less than 25 mV). This is achieved by the judicious use of an electrolyte.

A comparable situation is observed when a negative ion such as PO_4^{3-} is added to a suspension of positively charged particles such as bismuth subnitrate. Ionic and nonionic surfactants and lyophilic polymers also have been used to flocculate particles in suspension. Polymers, which act by forming a "bridge" between particles, may be the most efficient additives for inducing flocculation. Thus, it has been shown that the sedimentation volume is higher in suspensions flocculated with an anionic heteropolysaccharide than when electrolytes were used.

Work by Matthews and Rhodes,[51-53] involving both experimental and theoretical studies, has confirmed the formulation principles proposed by Martin and Haines. The suspensions used by Matthews and Rhodes contained 2.5% w/v of griseofulvin as a fine powder together with the anionic surfactant sodium dioxyethylated dodecyl sulfate (10^{-3} molar) as a wetting agent. Increasing concentrations of aluminum chloride were added and the sedimentation height (equivalent to the sedimentation volume, see page 295) and the zeta potential recorded. Flocculation occurred when a concentration of 10^{-3} molar aluminum chloride was reached. At this point the zeta potential had fallen from −46.4 mV to −17.0 mV. Further reduction of the zeta potential, to −4.5 mV by use of 10^{-2} molar aluminum chloride did not increase sedimentation height, in agreement with the principles shown in Fig 19-35.

Matthews and Rhodes then went on to show, by computer analysis, that the DLVO theory (see page 285) predicted the results obtained, namely, that the griseofulvin suspensions under investigation would remain deflocculated when the concentration of aluminum chloride was 10^{-4} molar or less. Only at concentrations in the range of 10^{-3} to 10^{-2} molar aluminum chloride did the theoretical plots show deep primary minima, indicative of flocculation. These occurred at a distance of separation between particles of approximately

50 Å, and led Matthews and Rhodes to conclude that coagulation had taken place in the primary minimum.

Schneider, *et al*[54] have published details of a laboratory investigation (suitable for undergraduates) that combines calculations based on the DLVO theory carried out with an interactive computer program with actual sedimentation experiments performed on simple systems.

Flocculation in Structured Vehicles—The ideal formulation for a suspension would seem to be when flocculated particles are supported in a structured vehicle.

As shown in Fig 19-34 (under C), the process involves dispersion of the particles and their subsequent flocculation. Finally, a lyophilic polymer is added to form the structured vehicle. In developing the formulation, care must be taken to ensure the absence of any incompatibility between the flocculating agent and the polymer used for the structured vehicle. A limitation is that virtually all the structured vehicles in common use are hydrophilic colloids and carry a negative charge. This means that an incompatibility arises if the charge on the particles is originally negative. Flocculation in this instance requires the addition of a positively charged flocculating agent or ion; in the presence of such a material, the negatively charged suspending agent may coagulate and lose its suspendability. This situation does not arise with particles that bear a positive charge, as the negative flocculating agent which the formulator must employ is compatible with the similarly charged suspending agent.

Chemical Stability of Suspensions—Particles that are completely insoluble in a liquid vehicle are unlikely to undergo most chemical reactions leading to degradation. However, most drugs in suspension have a finite solubility, even though this may be of the order of fractions of a microgram per mL. As a result, the material in solution may be susceptible to degradation. However, Tingstad and coworkers[55] developed a simplified method for determining the stability of drugs in suspension. The approach is based on the assumptions that (1) degradation takes place only in the solution and is first order, (2) the effect of temperature on drug solubility and reaction rate conforms with classical theory, and (3) dissolution is not rate-limiting on degradation.

Preparation of Suspensions—The small-scale preparation of suspensions may be readily undertaken by the practicing pharmacist with the minimum of equipment. The initial dispersion of the particles is best carried out by trituration in a mortar, the wetting agent being added in small increments to the powder. Once the particles have been wetted adequately, the slurry may be transferred to the final container. The next step depends on whether the deflocculated particles are to be suspended in a structured vehicle, flocculated, or flocculated and then suspended. Regardless of which of the alternative procedures outlined in Fig 19-34 is employed, the various manipulations can be carried out easily in the bottle, especially if an aqueous solution of the suspending agent has been prepared beforehand.

For a detailed discussion of the methods used in the large-scale production of suspensions, see the relevant section in Chapter 82.

Emulsions in Pharmacy

An emulsion is a dispersed system containing at least two immiscible liquid phases. The majority of conventional emulsions in pharmaceutical use have dispersed particles ranging in diameter from 0.1 to 100 μm. As with suspensions, emulsions are thermodynamically unstable as a result of the excess free energy associated with the surface of the droplets. The dispersed droplets, therefore, strive to come together and reduce the surface area. In addition to this flocculation effect, also observed with suspensions, the dispersed particles can coalesce, or fuse, and this can result in the eventual destruction of the emulsion. In order to minimize this effect a third component, the *emulsifying agent*, is added to the system to improve its stability. The choice of emulsifying agent is critical to the preparation of an emulsion possessing optimum stability. The efficiency of present-day emulsifiers permits the preparation of emulsions which are stable for many months and even years, even though they are thermodynamically unstable.

Emulsions are widely used in pharmacy and medicine, and emulsified materials can possess advantages not observed when formulated in other dosage forms. Thus, certain medicinal agents having an objectionable taste have been made more palatable for oral administration when formulated in an emulsion. The principles of emulsification have been applied extensively in the formulation of dermatological creams and lotions. Intravenous emulsions of contrast media have been developed to assist the physician in undertaking X-ray examinations of the body organs while exposing the patient to the minimum of radiation. Considerable attention has been directed towards the use of sterile, stable intravenous emulsions containing fat, carbohydrate, and vitamins all in one preparation. Such products are administered to patients unable to assimilate these vital materials by the normal oral route.

Emulsions offer potential in the design of systems capable of giving controlled rates of drug release and of affording protection to drugs susceptible to oxidation or hydrolysis. There is still a need for well-characterized dermatological products with reproducible properties, regardless of whether these products are antibacterial, sustained-release, protective, or emollient lotions, creams or ointments. The principle of emulsification is involved in an increasing number of aerosol products.

The pharmacist must be familiar with the types of emulsions and the properties and theories underlying their preparation and stability; such is the purpose of the remainder of this chapter. Microemulsions, which can be regarded as isotropic, swollen micellar systems are discussed in Chapter 83.

Emulsion Type and Means of Detection

A stable emulsion must contain at least three components; namely, the dispersed phase, the dispersion medium, and the emulsifying agent. Invariably, one of the two immiscible liquids is aqueous while the second is an oil. Whether the aqueous or the oil phase becomes the dispersed phase depends primarily on the emulsifying agent used and the relative amounts of the two liquid phases. Hence, an emulsion in which the oil is dispersed as droplets throughout the aqueous phase is termed an oil-in-water, O/W, emulsion. When water is the dispersed phase and an oil the dispersion medium, the emulsion is of the water-in-oil, W/O, type. Most pharmaceutical emulsions designed for oral administration are of the O/W type; emulsified lotions and creams are either O/W or W/O, depending on their use. Butter and salad creams are W/O emulsions.

Recently, so-called *multiple* emulsions have been developed with a view to delaying the release of an active ingredient. In these types of emulsions three phases are present, ie, the emulsion has the form W/O/W or O/W/O. In these

"emulsions within emulsions," any drug present in the innermost phase must now cross two phase boundaries to reach the external, continuous, phase.

It is important for the pharmacist to know the type of emulsion he has prepared or is dealing with, since this can affect its properties and performance. Unfortunately, the several methods available can give incorrect results, and so the type of emulsion determined by one method should always be confirmed by means of a second method.

Dilution Test—This method depends on the fact that an O/W emulsion can be diluted with water and a W/O emulsion with oil. When oil is added to an O/W emulsion or water to a W/O emulsion, the additive is not incorporated into the emulsion and separation is apparent. The test is greatly improved if the addition of the water or oil is observed microscopically.

Conductivity Test—An emulsion in which the continuous phase is aqueous can be expected to possess a much higher conductivity than an emulsion in which the continuous phase is an oil. Accordingly, it frequently happens that when a pair of electrodes, connected to a lamp and an electrical source, are dipped into an O/W emulsion, the lamp lights due to passage of a current between the two electrodes. If the lamp does not light, it is assumed that the system is W/O.

Dye-Solubility Test—The knowledge that a water-soluble dye will dissolve in the aqueous phase of an emulsion while an oil-soluble dye will be taken up by the oil phase provides a third means of determining emulsion type. Thus, if microscopic examination shows that a water-soluble dye has been taken up by the continuous phase, we are dealing with an O/W emulsion. If the dye has not stained the continuous phase, the test is repeated using a small amount of an oil-soluble dye. Coloring of the continuous phase confirms that the emulsion is of the W/O type.

Formation and Breakdown of Dispersed Liquid Droplets

An emulsion exists as the result of two competing processes, namely, the dispersion of one liquid throughout another as droplets, and the combination of these droplets to reform the initial bulk liquids. The first process increases the free energy of the system, while the second works to reduce the free energy. Accordingly, the second process is spontaneous and continues until breakdown is complete; ie, the bulk phases are reformed.

It is of little use to form a well-dispersed emulsion if it quickly breaks down. Similarly, unless adequate attention is given to achieving an optimum dispersion during preparation, the stability of an emulsion system may be compromised from the start. Dispersion is brought about by well-designed and well-operated machinery, capable of producing droplets in a relatively short period of time. Such equipment is discussed in Chapter 83. The reversal back to the bulk phases is minimized by utilizing those parameters which influence the stability of the emulsion once it is formed.

Dispersion Process To Form Droplets—Consider two immiscible liquid phases in a test tube. In order to disperse one liquid as droplets within the other, the interface between the two liquids must be disturbed and expanded to a sufficient degree so that "fingers" or threads of one liquid pass into the second liquid, and *vice versa*. These threads are unstable, and become varicosed or beaded. The beads separate and become spherical, as illustrated in Fig 19-36. Depending on the agitation or the shear rate used, larger droplets are also deformed to give small threads, which in turn produce smaller drops.

The time of agitation is important. Thus, the mean size of droplets decreases rapidly in the first few seconds of agitation. The limiting size range is generally reached within 1 to 5 minutes, and results from the number of droplets coalescing being equivalent to the number of new droplets being formed. It is uneconomical to continue agitation any further.

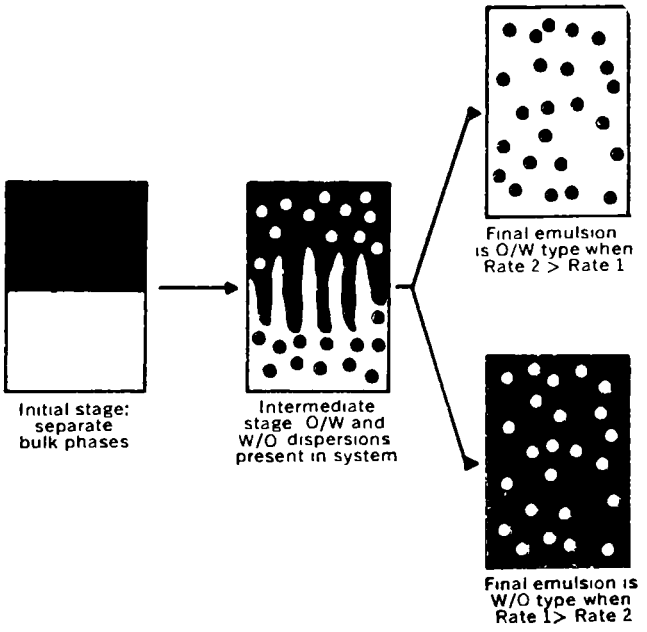


Fig 19-36. Effect of rate of coalescence on emulsion type. Rate 1: O/W coalescence rate; Rate 2: W/O coalescence rate. ●: oil; O: water. For an explanation of Rates 1 and 2, refer to the discussion of Davies on page 304.

The liquids may be agitated or sheared by several means. Shaking is commonly employed, especially when the components are of low viscosity. Intermittent shaking is frequently more efficient than continual shaking, possibly because the short time interval between shakes allows the thread which is forced across the interface time to break down into drops which are then isolated in the opposite phase. Continuous, rapid agitation tends to hinder this breakdown to form drops. A mortar and pestle is employed frequently in the extemporaneous preparation of emulsions. It is not a very efficient technique and is not used on a large scale. Improved dispersions are achieved by the use of high-speed mixers, blenders, colloid mills and homogenizers. Ultrasonic techniques also have been employed and are described in Chapter 83.

The phenomenon of spontaneous emulsification, as the name implies, occurs without any external agitation. There is, however, an internal agitation arising from certain physicochemical processes that affect the interface between the two bulk liquids. For a description of this process, see Davies and Rideal in the *Bibliography*.

Coalescence of Droplets—Coalescence is a process distinct from flocculation (aggregation), which commonly precedes it. While flocculation is the clumping together of particles, coalescence is the fusing of the agglomerates into a larger drop, or drops. Coalescence is usually rapid when two immiscible liquids are shaken together, since there is no large energy barrier to prevent fusion of drops and reformation of the original bulk phases. When an emulsifying agent is added to the system, flocculation still may occur but coalescence is reduced to an extent depending on the efficacy of the emulsifying agent to form a stable, coherent interfacial film. It is therefore possible to prepare emulsions that are flocculated, yet which do not coalesce. In addition to the interfacial film around the droplets acting as a mechanical

barrier, the drops also are prevented from coalescing by the presence of a thin layer of continuous phase between particles clumped together.

Davies[56] showed the importance of coalescence rates in determining emulsion type; this work is discussed in more detail on page 304.

Emulsifying Agent

The process of coalescence can be reduced to insignificant levels by the addition of a third component—the emulsifying agent or emulsifier. The choice of emulsifying agent is frequently critical in developing a successful emulsion, and the pharmacist should be aware of

The desirable properties of emulsifying agents.
How different emulsifiers act to optimize emulsion stability.
How the type and physical properties of the emulsion can be affected by the emulsifying agent.

Desirable Properties

Some of the desirable properties of an emulsifying agent are that it should

1. Be surface-active and reduce surface tension to below 10 dynes/cm.
2. Be adsorbed quickly around the dispersed drops as a condensed, nonadherent film which will prevent coalescence.
3. Impart to the droplets an adequate electrical potential so that mutual repulsion occurs.
4. Increase the viscosity of the emulsion.
5. Be effective in a reasonably low concentration.

Not all emulsifying agents possess these properties to the same degree; in fact, not every good emulsifier necessarily possesses all these properties. Further, there is no one "ideal" emulsifying agent because the desirable properties of an emulsifier depend, in part, on the properties of the two immiscible phases in the particular system under consideration.

Interfacial Tension—Lowering of interfacial tension is one way in which the increased surface free energy associated with the formation of droplets, and hence surface area, in an emulsion can be reduced (Eq 29). Assuming the droplets to be spherical, it can be shown that

$$\Delta F = \frac{6\gamma V}{d} \qquad (38)$$

where V is the volume of dispersed phase in mL and d is the mean diameter of the particles. In order to disperse 100 mL of oil as 1-μm (10^{-4}-cm) droplets in water when $\gamma_{O/W} = 50$ dynes/cm, requires an energy input of

$$\Delta F = \frac{6 \times 50 \times 100}{1 \times 10^{-4}} = 30 \times 10^7 \text{ ergs}$$

$$= 30 \text{ joules or } 30/4.184 = 7.2 \text{ cal}$$

In the above example the addition of an emulsifier that will reduce γ from 50 to 5 dynes/cm will reduce the surface free energy from 7.2 to around 0.7 cal. Likewise, if the interfacial tension is reduced to 0.5 dyne/cm, a common occurrence, the original surface free energy is reduced a hundredfold. Such a reduction can help to maintain the surface area generated during the dispersion process.

Film Formation—The major requirement of a potential emulsifying agent is that it readily form a film around each droplet of dispersed material. The main purpose of this film—which can be a monolayer, a multilayer, or a collection of small particles adsorbed at the interface—is to form a barrier which prevents the coalescence of droplets that come into contact with one another. For the film to be an efficient barrier, it should possess some degree of surface elasticity and should not thin out and rupture when sandwiched between two droplets. If broken, the film should have the capacity to reform rapidly.

Electrical Potential—The origin of an electrical potential at the surface of a droplet has been discussed earlier in the chapter. Insofar as emulsions are concerned, the presence of a well-developed charge on the droplet surface is significant in promoting stability by causing repulsion between approaching drops. This potential is likely to be greater when an ionized emulsifying agent is employed.

Concentration of Emulsifier—The main objective of an emulsifying agent is to form a condensed film around the droplets of the dispersed phase. An inadequate concentration will do little to prevent coalescence. Increasing the emulsifier concentration above an optimum level achieves little in terms of increased stability. In practice the aim is to use the minimum amount consistent with producing a satisfactory emulsion.

It frequently helps to have some idea of the amount of emulsifier required to form a condensed film, one molecule thick, around each droplet. Suppose we wish to emulsify 50 g of an oil, density = 1.0, in 50 g of water. The desired particle diameter is 1 μm. Thus,

$$\text{Particle diameter} = 1\ \mu\text{m} = 1 \times 10^{-4}\ \text{cm}$$

$$\text{Volume of particle} = \frac{\pi d^3}{6} = 0.524 \times 10^{-12}\ \text{cm}^3$$

$$\text{Total number of particles in 50 g} = \frac{50}{0.524 \times 10^{-12}} = 95.5 \times 10^{12}$$

$$\text{Surface area of each particle} = \pi d^2 = 3.142 \times 10^{-8}\ \text{cm}^2$$

$$\text{Total surface area} = 3.142 \times 10^{-8} \times 95.5 \times 10^{12} = 300 \times 10^4\ \text{cm}^2$$

If the area each molecule occupies at the oil/water interface is 30 Å^2 (30×10^{-16} cm^2), we require

$$\frac{300 \times 10^4}{30 \times 10^{16}} = 1 \times 10^{21} \text{ molecules}$$

A typical emulsifying agent might have a molecular weight of 1000. Thus, the required weight is

$$\frac{1000 \times 10^{21}}{6.023 \times 10^{23}} = 1.66 \text{ g}$$

To emulsify 10 g of oil would require 0.33 g of the emulsifying agent, etc. While the approach is an oversimplification of the problem, it does at least allow the formulator to make a reasonable estimate of the required concentration of emulsifier.

Emulsion Rheology—The emulsifying agent and other components of an emulsion can affect the rheologic behavior of an emulsion in several ways and these are summarized in Table XVI. It should be borne in mind that the droplets of the internal phase are deformable under shear and that the adsorbed layer of emulsifier affects the interactions between adjacent droplets and also between a droplet and the continuous phase.

The means by which the rheological behavior of emulsions can be controlled have been discussed by Rogers.[58]

Mechanism of Action

Emulsifying agents may be classified in accordance with the type of film they form at the interface between the two phases.

Monomolecular Films—Those surface-active agents which are capable of stabilizing an emulsion do so by form-

Table XVI—Factors Influencing Emulsion Viscosity[57]

1. Internal phase
   a. Volume concentration (ϕ); hydrodynamic interaction between globules; flocculation, leading to formation of globule aggregates.
   b. Viscosity (η_1); deformation of globules in shear.
   c. Globule size, and size distribution, technique used to prepare emulsion; interfacial tension between the two liquid phases: globule behavior in shear; interaction with continuous phase; globule interaction.
   d. Chemical constitution.
2. Continuous phase
   a. Viscosity (η_0), and other rheological properties.
   b. Chemical constitution, polarity, pH; potential energy of interaction between globules.
   c. Electrolyte concentration if polar medium.
3. Emulsifying agent
   a. Chemical constitution; potential energy of interaction between globules.
   b. Concentration, and solubility in internal and continuous phases; emulsion type; emulsion inversion; solubilization of liquid phases in micelles.
   c. Thickness of film adsorbed around globules, and its rheological properties, deformation of globules in shear; fluid circulation within globules.
   d. Electroviscous effect.
4. Additional stabilizing agents
   Pigments, hydrocolloids, hydrous oxides; effect on rheologic properties of liquid phases, and interfacial boundary region.

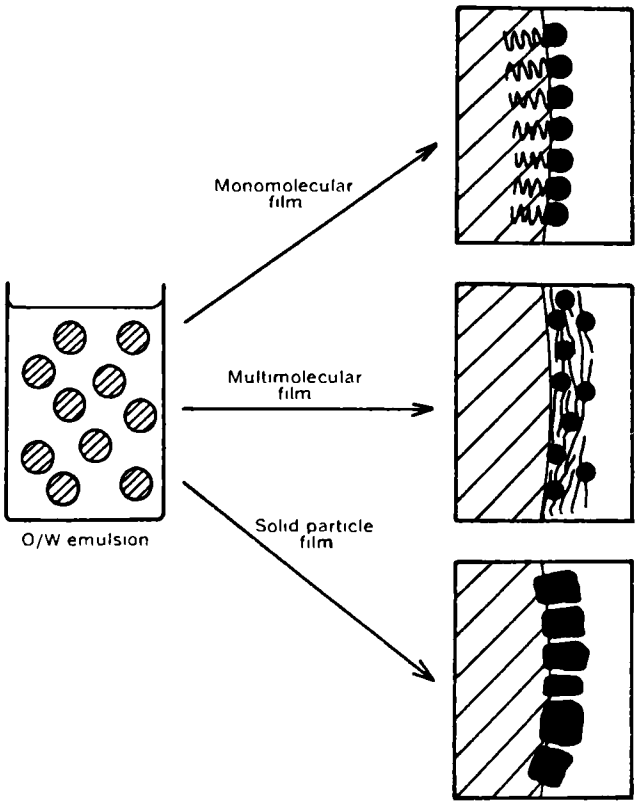


Fig 19-37. Types of films formed by emulsifying agents at the oil/water interface. Orientations are shown for O/W emulsions. ▨: oil; □: water.

ing a monolayer of adsorbed molecules or ions at the oil/water interface (Fig 19-37). In accordance with Gibbs' law (Eq 29) the presence of an interfacial excess necessitates a reduction in interfacial tension. This results in a more stable emulsion because of a proportional reduction in the surface free energy. Of itself, this reduction is probably not the main factor promoting stability. More significant is the fact that the droplets are surrounded now by a coherent monolayer which prevents coalescence between approaching droplets. If the emulsifier forming the monolayer is ionized, the presence of strongly charged and mutually repelling droplets increases the stability of the system. With unionized, nonionic surface-active agents, the particles may still carry a charge; this arises from adsorption of a specific ion or ions from solution.

Multimolecular Films—Hydrated lyophilic colloids form multimolecular films around droplets of dispersed oil (Fig 19-37). The use of these agents has declined in recent years because of the large number of synthetic surface-active agents available which possess well-marked emulsifying properties. While these hydrophilic colloids are adsorbed at an interface (and can be regarded therefore as "surface-active"), they do not cause an appreciable lowering in surface tension. Rather, their efficiency depends on their ability to form strong, coherent multimolecular films. These act as a coating around the droplets and render them highly resistant to coalescence, even in the absence of a well-developed surface potential. Furthermore, any hydrocolloid not adsorbed at the interface increases the viscosity of the continuous aqueous phase; this enhances emulsion stability.

Solid Particle Films—Small solid particles that are wetted to some degree by both aqueous and nonaqueous liquid phases act as emulsifying agents. If the particles are too hydrophilic, they remain in the aqueous phase; if too hydrophobic, they are dispersed completely in the oil phase. A second requirement is that the particles are small in relation to the droplets of the dispersed phase (Fig 19-37).

Chemical Types

Emulsifying agents may also be classified in terms of their chemical structure; there is some correlation between this classification and that based on the mechanism of action. For example, the majority of emulsifiers forming monomolecular films are synthetic, organic materials. Most of the emulsifiers that form multimolecular films are obtained from natural sources and are organic. A third group is composed of solid particles, invariably inorganic, that form films composed of finely divided solid particles.

Accordingly, the classification adopted divides emulsifying agents into *synthetic*, *natural*, and *finely dispersed solids* (Table XVII). A fourth group, the *auxiliary materials* (Table XVIII), are weak emulsifiers. The agents listed are designed to illustrate the various types available; they are not meant to be exhaustive.

Synthetic Emulsifying Agents—This group of surface-active agents which act as emulsifiers may be subdivided into anionic, cationic, and nonionic, depending on the charge possessed by the surfactant.

Anionics—In this subgroup the surfactant ion bears a negative charge. The potassium, sodium, and ammonium salts of lauric and oleic acid are soluble in water and are good O/W emulsifying agents. They do, however, have a disagreeable taste and are irritating to the gastrointestinal tract; this limits them to emulsions prepared for external use. Potassium laurate, a typical example, has the structure

$$CH_3(CH_2)_{10}COO^- \quad K^+$$

Solutions of alkali soaps have a high pH; they start to precipitate out of solution below pH 10 because the unionized fatty acid is now formed, and this has a low aqueous solubility. Further, the free fatty acid is ineffective as an emulsifier and so emulsions formed from alkali soaps are not stable at pH values less than about 10.

The calcium, magnesium and aluminum salts of fatty acids, often termed the metallic soaps, are water insoluble and result in W/O emulsions.

Table XVII—Classification of Emulsifying Agents

Type	Type of film	Examples
Synthetic (surface-active agents)	Monomolecular	*Anionic* Soaps Potassium laurate Triethanolamine stearate Sulfates Sodium lauryl sulfate Alkyl polyoxyethylene sulfates Sulfonates Dioctyl sodium sulfosuccinate *Cationic:* Quaternary ammonium compounds Cetyltrimethylammonium bromide Lauryldimethylbenzylammonium chloride *Nonionic.* Polyoxyethylene fatty alcohol ethers Sorbitan fatty acid esters Polyoxyethylene sorbitan fatty acid esters
Natural	Multimolecular	*Hydrophilic colloids.* Acacia Gelatin
	Monomolecular	Lecithin Cholesterol
Finely divided solids	Solid particle	*Colloidal clays* Bentonite Veegum *Metallic hydroxides:* Magnesium hydroxide

Table XVIII—Auxiliary Emulsifying Agents[55]

Product	Source and composition	Principal use
Bentonite	Colloidal hydrated aluminum silicate	Hydrophilic thickening agent and stabilizer for O/W and W/O lotions and creams
Cetyl alcohol	Chiefly $C_{16}H_{33}OH$	Lipophilic thickening agent and stabilizer for O/W lotions and ointments
Glyceryl monostearate	$C_{17}H_{35}COOCH_2CHOHCH_2OH$	Lipophilic thickening agent and stabilizer for O/W lotions and ointments
Methylcellulose	Series of methyl esters of cellulose	Hydrophilic thickening agent and stabilizer for O/W emulsions; weak O/W emulsifier
Sodium alginate	The sodium salt of alginic acid, a purified carbohydrate extracted from giant kelp	Hydrophilic thickening agent and stabilizer for O/W emulsions
Sodium carboxymethylcellulose	Sodium salt of the carboxymethyl esters of cellulose	Hydrophilic thickening agent and stabilizer for O/W emulsions
Stearic acid	A mixture of solid acids from fats, chiefly stearic and palmitic	Lipophilic thickening agent and stabilizer for O/W lotions and ointments. Forms a true emulsifier when reacted with an alkali
Stearyl alcohol	Chiefly $C_{18}H_{37}OH$	Lipophilic thickening agent and stabilizer for O/W lotions and ointments
Veegum	Colloidal magnesium aluminum silicate	Hydrophilic thickening agent and stabilizer for O/W lotions and creams

Another class of soaps are salts formed from a fatty acid and an organic amine such as triethanolamine. While these O/W emulsifiers are also limited to external preparations, their alkalinity is considerably less than that of the alkali soaps and they are active as emulsifiers down to around pH 8. These agents are less irritating than the alkali soaps.

Sulfated alcohols are neutralized sulfuric acid esters of such fatty alcohols as lauryl and cetyl alcohol. These compounds are an important group of pharmaceutical surfactants. They are used chiefly as wetting agents, although they do have some value as emulsifiers, particularly, when used in conjunction with an auxiliary agent. A frequently used compound is sodium lauryl sulfate.

$$CH_3(CH_2)_{10}CH_2OSO_3^-\ \ Na^+$$

Sulfonates are a class of compounds in which the sulfur atom is connected directly to the carbon atom, giving the general formula

$$CH_3(CH_2)_nCH_2SO_3^-\ \ Na^+$$

Sulfonates have a higher tolerance to calcium ions and do not hydrolyze as readily as the sulfates. A widely used surfactant of this type is dioctyl sodium sulfosuccinate.

Cationics—The surface activity in this group resides in the positively charged cation. These compounds have marked bactericidal properties. This makes them desirable in emulsified anti-infective products such as skin lotions and creams. The pH of an emulsion prepared with a cationic emulsifier lies in the pH 4–6 range. Since this includes the normal pH of the skin, cationic emulsifiers are advantageous in this regard also.

Cationic agents are weak emulsifiers and are generally formulated with a stabilizing or auxiliary emulsifying agent such as cetostearyl alcohol. The only group of cationic agents used extensively as emulsifying agents are the quaternary ammonium compounds. An example is cetyltrimethylammonium bromide.

$$CH_3(CH_2)_{14}CH_2N^+(CH_3)_3\ \ Br^-$$

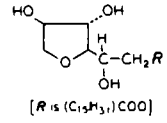
Cationic emulsifiers should not be used in the same formulation with anionic emulsifiers as they will interact. While the incompatibility may not be immediately apparent as a precipitate, virtually all of the desired antibacterial activity will generally have been lost.

Nonionics—These undissociated surfactants find widespread use as emulsifying agents when they possess the proper balance of hydrophilic and lipophilic groups within the molecule. Their popularity is based on the fact that, unlike the anionic and cationic types, nonionic emulsifiers are not susceptible to pH changes and the presence of electrolytes. The number of nonionic agents available is legion; the most frequently used are the glyceryl esters, polyoxyethylene glycol esters and ethers, and the sorbitan fatty acid esters and their polyoxyethylene derivatives.

A glyceryl ester, such as glyceryl monostearate, is too lipophilic to serve as a good emulsifier; it is widely used as an auxiliary agent (Table XVIII) and has the structure

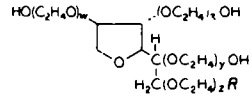
$$\begin{array}{l} CH_2OOCC_{17}H_{35} \\ | \\ CHOH \\ | \\ CH_2OH \end{array}$$

Sorbitan fatty acid esters, such as sorbitan monopalmitate



[R is $(C_{15}H_{31})COO$]

are nonionic oil-soluble emulsifiers that promote W/O emulsions. The polyoxyethylene sorbitan fatty acid esters, such as polyoxyethylene sorbitan monopalmitate, are hydrophilic water-soluble derivatives that favor O/W emulsions.



[Sum of w, x, y and z is 20, R is $(C_{15}H_{31})COO$]

Polyoxyethylene glycol esters, such as the monostearate, $C_{17}H_{35}COO(CH_2OCH_2)_nH$, also are used widely.

Very frequently, the best results are obtained from blends of nonionic emulsifiers. Thus, an O/W emulsifier customarily will be used in an emulsion with a W/O emulsifier. When blended properly, the nonionics produce fine-textured stable emulsions.

Natural Emulsifying Agents—Of the numerous emulsifying agents derived from natural (ie, plant and animal) sources, consideration will be given only to acacia, gelatin, lecithin, and cholesterol. Many other natural materials are only sufficiently active to function as auxiliary emulsifying agents or stabilizers.

Acacia is a carbohydrate gum that is soluble in water and forms O/W emulsions. Emulsions prepared with acacia are stable over a wide pH range. Because it is a carbohydrate it is necessary to preserve acacia emulsions against microbial attack by the use of a suitable preservative. The gum can be precipitated from aqueous solution by the addition of high concentrations of electrolytes or solvents less polar than water, such as alcohol.

Gelatin, a protein, has been used for many years as an emulsifying agent. Gelatin can have two isoelectric points, depending on the method of preparation. So-called Type A gelatin, derived from an acid-treated precursor, has an isoelectric point of between pH 7 and 9. Type B gelatin, obtained from an alkali-treated precursor, has an isoelectric point of approximately pH 5. Type A gelatin acts best as an emulsifier around pH 3, where it is positively charged; on the other hand, Type B gelatin is best used around pH 8, where it is negatively charged. The question as to whether the gelatin is positively or negatively charged is fundamental to the stability of the emulsion when other charged emulsifying agents are present. In order to avoid an incompatibility, all emulsifying agents should carry the same sign. Thus, if gums (such as tragacanth, acacia or agar) which are negatively charged are to be used with gelatin, Type B material should be used at an alkaline pH. Under these conditions the gelatin is similarly negatively charged.

Lecithin is a phospholipid which, because of its strongly hydrophilic nature, produces O/W emulsions. It is liable to microbial attack and tends to darken on storage.

Cholesterol is a major constituent of wool alcohols, obtained by the saponification and fractionation of wool fat. It is cholesterol that gives wool fat its capacity to absorb water and form a W/O emulsion.

Finely Dispersed Solids—This group of emulsifiers forms particulate films around the dispersed droplets and produces emulsions which, while coarse-grained, have considerable physical stability. It appears possible that any solid can act as an emulsifying agent of this type, provided it is reduced to a sufficiently fine powder. In practice the group of compounds used most frequently are the colloidal clays.

Several colloidal clays find application in pharmaceutical emulsions; the most frequently used are bentonite, a colloidal aluminum silicate, and Veegum (*Vanderbilt*), a colloidal magnesium aluminum silicate.

Bentonite is a white to gray, odorless, and tasteless powder that swells in the presence of water to form a translucent suspension with a pH of about 9. Depending on the sequence of mixing it is possible to prepare both O/W and W/O emulsions. When an O/W emulsion is desired, the bentonite is first dispersed in water and allowed to hydrate so as to form a magma. The oil phase is then added gradually with constant trituration. Since the aqueous phase is always in excess, the O/W emulsion type is favored. To prepare a W/O emulsion, the bentonite is first dispersed in oil; the water is then added gradually.

While Veegum is used as a solid particle emulsifying agent, it is employed most extensively as a stabilizer in cosmetic lotions and creams. Concentrations of less than 1% Veegum will stabilize an emulsion containing anionic or nonionic emulsifying agents.

Auxiliary Emulsifying Agents—Included under this heading are those compounds which are normally incapable themselves of forming stable emulsions. Their main value lies in their ability to function as thickening agents and thereby help stabilize the emulsion. Agents in common use are listed in Table XVIII.

Emulsifying Agents and Emulsion Type

For a molecule, ion, colloid, or particle to be active as an emulsifying agent, it must have some affinity for the interface between the dispersed phase and the dispersion medium. With the mono- and multilayer films the emulsifier is in solution and, therefore, must be soluble to some extent in one or both of the phases. At the same time it must not be overly soluble in either phase, otherwise it will remain in the bulk of that phase and not be adsorbed at the interface. This balanced affinity for the two phases also must be evident with finely divided solid particles used as emulsifying agents. If their affinity, as evidenced by the degree to which they are wetted, is either predominantly hydrophilic or hydrophobic, they will not function as effective wetting agents.

The great majority of the work on the relation between

Table XIX—Approximate HLB Values for a Number of Emulsifying Agents

Generic or chemical name	HLB
Sorbitan trioleate	1.8
Sorbitan tristearate	2.1
Propylene glycol monostearate	3.4
Sorbitan sesquioleate	3.7
Glycerol monostearate (non self-emulsifying)	3.8
Sorbitan monooleate	4.3
Propylene glycol monolaurate	4.5
Sorbitan monostearate	4.7
Glyceryl monostearate (self-emulsifying)	5.5
Sorbitan monopalmitate	6.7
Sorbitan monolaurate	8.6
Polyoxyethylene-4-lauryl ether	9.5
Polyethylene glycol 400 monostearate	11.6
Polyoxyethylene-4-sorbitan monolaurate	13.3
Polyoxyethylene-20-sorbitan monooleate	15.0
Polyoxyethylene-20-sorbitan monopalmitate	15.6
Polyoxyethylene-20-sorbitan monolaurate	16.7
Polyoxyethylene-40-stearate	16.9
Sodium oleate	18.0
Sodium lauryl sulfate	40.0

emulsifier and emulsion type has been concerned with surface-active agents that form interfacial monolayers. The present discussion, therefore, will concentrate on this class of agents.

Hydrophile–Lipophile Balance—As the emulsifier becomes more hydrophilic, its solubility in water increases and the formation of an O/W emulsion is favored. Conversely, W/O emulsions are favored with the more lipophilic emulsifiers. This led to the concept that the type of emulsion is related to the balance between hydrophilic and lipophilic solution tendencies of the surface-active emulsifying agent.

Griffin[59] developed a scale based on the balance between these two opposing tendencies. This so-called *HLB scale* is a numerical scale, extending from 1 to approximately 50. The more hydrophilic surfactants have high HLB numbers (in excess of 10), while surfactants with HLB numbers from 1 to 10 are considered to be lipophilic. Surfactants with a proper balance in their hydrophilic and lipophilic affinities are effective emulsifying agents since they concentrate at the oil/water interface. The relationship between HLB values and the application of the surface-active agent is shown in Table XV. Some commonly used emulsifiers and their HLB numbers are listed in Table XIX. The utility of the HLB system in rationalizing the choice of emulsifying agents when formulating an emulsion will be discussed in a later section.

Rate of Coalescence and Emulsion Type—Davies[56] indicated that the type of emulsion produced in systems prepared by shaking is controlled by the relative coalescence rates of oil droplets dispersed in the oil. Thus, when a mixture of oil and water is shaken together with an emulsifying agent, a multiple dispersion is produced initially which contains oil dispersed in water and water dispersed in oil (Fig 19-36). The type of the final emulsion which results depends on whether the water or the oil droplets coalesce more rapidly. If the O/W coalescence rate (Rate 1) is much greater than W/O coalescence rate (Rate 2), a W/O emulsion is formed since the dispersed water droplets are more stable than the dispersed oil droplets. Conversely, if Rate 2 is significantly faster than Rate 1, the final emulsion is an O/W dispersion because the oil droplets are more stable.

According to Davies, the rate at which oil globules coalesce when dispersed in water is given by the expression

$$\text{Rate 1} = C_1 e^{-W_1/RT} \qquad (39)$$

The term C_1 is a collision factor which is directly proportional to the phase volume of the oil relative to the water, and is an inverse function of the viscosity of the continuous phase (water). W_1 defines an energy barrier made up of several contributing factors that must be overcome before coalescence can take place. First, it depends on the electrical potential of the dispersed oil droplets, since this affects repulsion. Second, with an O/W emulsion, the hydrated layer surrounding the polar portion of emulsifying agent must be broken down before coalescence can occur. This hydrated layer is probably around 10 Å thick with a consistency of butter. Finally, the total energy barrier depends on the fraction of the interface covered by the emulsifying agent.

Equation 40 describes the rate of coalescence of water globules dispersed in oil, namely

$$\text{Rate 2} = C_2 e^{-W_2/RT} \qquad (40)$$

Here, the collision factor C_2 is a function of the water/oil phase volume ratio divided by the viscosity of the oil phase. The energy barrier W_2 is, as before, related to the fraction of the interface covered by the surface-active agent. Another contributing factor is the number of —CH_2— groups in the emulsifying agent; the longer the alkyl chain of the emulsifier, the greater the gap that has to be bridged if one water droplet is to combine with a second drop.

Davies[56] showed that the HLB concept is related to the distribution characteristics of the emulsifying agent between the two immiscible phases. An emulsifier with an HLB of less than 7 will be preferentially soluble in the oil phase and will favor formation of a W/O emulsion. Surfactants with an HLB value in excess of 7 will be distributed in favor of the aqueous phase and will promote O/W emulsions.

Preparation of Emulsions

Several factors must be taken into account in the successful preparation and formulation of emulsified products. Usually, the type of emulsion (ie, O/W or W/O) is specified; if not, it probably will be implied from the anticipated use of the product. The formulator's attention is focused primarily on the selection of the emulsifying agent, or agents, necessary to achieve a satisfactory product. No incompatibilities should occur between the various emulsifiers and the several components commonly present in pharmaceutical emulsions. Finally, the product should be prepared in such a way as not to prejudice the formulation.

Selection of Emulsifying Agents

The selection of the emulsifying agent, or agents, is of prime importance in the successful formulation of an emulsion. In addition to its emulsifying properties, the pharmacist must ensure that the material chosen is nontoxic and that the taste, odor, and chemical stability are compatible with the product. Thus, an emulsifying agent which is entirely suitable for inclusion in a skin cream may be unacceptable in the formulation of an oral preparation due to its potential toxicity. This consideration is most important when formulating intravenous emulsions.

The HLB System—With the increasing number of available emulsifiers, particularly the nonionics, the selection of emulsifiers for a product was essentially a trial-and-error procedure. Fortunately, the work of Griffin[59,60] provided a logical means of selecting emulsifying agents. Griffin's method, based on the balance between the hydrophilic and lipophilic portions of the emulsifying agent, is now widely used and has come to be known as the *HLB system*. It is used most in the rational selection of combinations of non-

Table XX—Relationship between HLB Range and Surfactant Application

HLB range	Use
0–3	Antifoaming agents
4–6	W/O emulsifying agents
7–9	Wetting agents
8–18	O/W emulsifying agents
13–15	Detergents
10–18	Solubilizing agents

Table XXI—Required HLB Values for Some Common Emulsion Ingredients

Substance	W/O	O/W
Acid, stearic	...	17
Alcohol, cetyl	...	13
Lanolin, anhydrous	8	15
Oil, cottonseed	...	7.5
mineral oil, light	4	10–12
mineral oil, heavy	4	10.5
Wax, beeswax	5	10–16
microcrystalline	...	9.5
paraffin	...	9

ionic emulsifiers, and we shall limit our discussion accordingly.

As shown in Table XX, if an O/W emulsion is required, the formulator should use emulsifiers with an HLB in the range of 8–18. Emulsifiers with HLB values in the range of 4–6 are given consideration when a W/O emulsion is desired. Some typical examples are given in Table XIX.

Another factor is the presence or absence of any polarity in the material being emulsified, since this will affect the polarity required in the emulsifier. Again, as a result of extensive experimentation, Griffin evolved a series of "required HLB" values; ie, the HLB value required by a particular material if it is to be emulsified effectively. Some values for oils and related materials are contained in Table XXI. Naturally, the required HLB value differs depending on whether the final emulsion is O/W or W/O.

Fundamental to the utility of the HLB concept is the fact that the HLB values are algebraically additive. Thus, by using a low HLB surfactant with one having a high HLB it is possible to prepare blends having HLB values intermediate between those of the two individual emulsifiers. Naturally, one should not use emulsifiers that are incompatible. The following formula should serve as an example.

O/W Emulsion

Liquid petrolatum (Required HLB 10.5)	50 g
Emulsifying agents	5 g
Sorbitan monooleate (HLB 4.3)	
Polyoxyethylene 20 sorbitan monoleate (HLB 15.0)	
Water, qs	100 g

By simple algebra it can be shown that 4.5 parts by weight of sorbitan monooleate blended with 6.2 parts by weight of polyoxyethylene 20 sorbitan monooleate will result in a mixed emulsifying agent having the required HLB of 10.5. Since the formula calls for 5 g, the required weights are 2.1 g and 2.9 g, respectively. The oil-soluble sorbitan monooleate is dissolved in the oil and heated to 75°; the water-soluble polyoxyethylene 20 sorbitan monooleate is added to the aqueous phase which is heated to 70°. At this point the oil phase is mixed with the aqueous phase and the whole stirred continuously until cool.

The formulator is not restricted to these two agents to produce a blend with an HLB of 10.5. Table XXII shows

Table XXII—Nonionic Blends having HLB Values of 10.5

Surfactant blend	HLB	Required amounts (%) to give HLB = 10 5
Sorbitan tristearate	2.1	34.4
Polyoxethylene 20 sorbitan monostearate	14.9	65.6
Sorbitan monopalmitate	6.7	57.3
Polyoxyethylene 20 sorbitan monopalmitate	15.6	42.7
Sorbitan sesquioleate	3.7	48.5
Polyoxyethylene lauryl ether	16.9	51.5

the various proportions required, using other pairs of emulsifying agents, to form a blend of HLB 10.5. When carrying out preliminary investigations with a particular material to be emulsified, it is advisable to try several pairs of emulsifying agents. Based on an evaluation of the emulsions produced, it becomes possible to choose the best combination.

Occasionally, the required HLB of the oil may not be known, in which case it becomes necessary to determine this parameter. Various blends are prepared to give a wide range of HLB mixtures and emulsions are prepared in a standardized manner. The HLB of the blend used to emulsify the best product, selected on the basis of physical stability, is taken to be the required HLB of the oil. The experiment should be repeated using another combination of emulsifiers to confirm the value of the required HLB of the oil to within, say, ±1 HLB unit.

There are methods for finding the HLB value of a new surface-active agent. Griffin[60] developed simple equations which can be used to obtain an estimate with certain compounds. It has been shown that the ability of a compound to spread at a surface is related to its HLB. In another approach a linear relation between HLB and the logarithm of the dielectric constant for a number of nonionic surfactants has been observed. An interesting approach has been developed by Davies[56] and is related to his studies on the relative rates of coalescence of O/W and W/O emulsions (page 304). According to Davies, hydrophilic groups on the surfactant molecule make a positive contribution to the HLB number, whereas lipophilic groups exert a negative effect. Davies calculated these contributions and termed them HLB Group Numbers (Table XXIII). Provided the molecular structure of the surfactant is known, one simply adds the various group numbers in accordance with the following formula:

Table XXIII—HLB Group Numbers[61]

	Group number
Hydrophilic groups	
$-SO_4^-Na^+$	38.7
$-COO^-K^+$	21.1
$-COO^-Na^+$	19.1
N (tertiary amine)	9.4
Ester (sorbitan ring)	6.8
Ester (free)	2.4
—COOH	2.1
Hydroxyl (free)	1.9
—O—	1.3
Hydroxyl (sorbitan ring)	0.5
Lipophilic groups	
—CH—	
$-CH_2-$	
CH_3-	−0.475
=CH—	
Derived groups	
$-(CH_2-CH_2-O)-$	+0.33
$-(CH_2-CH_2-CH_2-O)-$	−0.15

$$HLB = \Sigma(\text{hydrophilic group numbers}) - m(\text{group number}/\text{—}CH_2\text{— group}) + 7$$

where m is the number of —CH_2— groups present in the surfactant. Poor agreement is found between the HLB values calculated by the use of group numbers and the HLB values obtained using the simple equations developed by Griffin. However, the student should realise that the absolute HLB values *per se* are of limited significance. The utility of the HLB approach (using values calculated by either Griffin's or Davies' equations) is to (i) provide the formulator with an idea of the relative balance of hydrophilicity and lipophilicity in a particular surfactant, and (ii) relate that surfactant's emulsifying and solubilizing properties to other surfactants. The formulator still needs to confirm experimentally that a particular formulation will produce a stable emulsion.

Later, Davies and Rideal[61] attempted to relate HLB to the C_{water}/C_{oil} partition coefficient and found good agreement for a series of sorbitan surfactants. Schott[62] showed, however, that the method does not apply to polyoxyethylated octylphenol surfactants. Schott concluded that "so far, the search for a universal correlation between HLB and another property of the surfactant which could be determined more readily than HLB has not been successful."

The HLB system gives no information as to the *amount* of emulsifier required. Having once determined the correct blend, the formulator must prepare another series of emulsions, all at the same HLB, but containing increasing concentrations of the emulsifier blend. Usually, the minimum concentration giving the desired degree of physical stability is chosen.

Mixed Emulsifying Agents—Emulsifying agents are frequently used in combination since a better emulsion usually is obtained. This enhancement may be due to several reasons, one or more of which may be operative in any one system. Thus, the use of a blend or mixture of emulsifiers may (1) produce the required hydrophile–lipophile balance in the emulsifier, (2) enhance the stability and cohesiveness of the interfacial film, and (3) affect the consistency and feel of the product.

The first point has been considered in detail in the previous discussion of the HLB system.

With regard to the second point, Schulman and Cockbain in 1940 showed that combinations of certain amphiphiles formed stable films at the air/water interface. It was postulated that the complex formed by these two materials (one, oil-soluble; the other, water-soluble) at the air/water interface was also present at the O/W interface. This interfacial complex was held to be responsible for the improved stability. For example, sodium cetyl sulfate, a moderately good O/W emulsifier, and elaidyl alcohol or cholesterol, both stabilizers for W/O emulsions, show evidence of an interaction at the air/water interface. Furthermore, an O/W emulsion prepared with sodium cetyl sulfate and elaidyl alcohol is much more stable than an emulsion prepared with sodium cetyl sulfate alone.

Elaidyl alcohol is the *trans* isomer. When oleyl alcohol, the *cis* isomer, is used with sodium cetyl sulfate, there is no evidence of complex formation at the air/water interface. Significantly, this combination does not produce a stable O/W emulsion either. Such a finding strongly suggests that a high degree of molecular alignment is necessary at the O/W interface to form a stable emulsion.

Finally, some materials are added primarily to increase the consistency of the emulsion. This may be done to increase stability or improve emolliency and feel. Examples include cetyl alcohol, stearic acid and beeswax.

When using combinations of emulsifiers, care must be taken to ensure their compatibility, as charged emulsifying agents of opposite sign are likely to interact and coagulate when mixed.

Small-Scale Preparation

Mortar and Pestle—This approach invariably is used only for those emulsions that are stabilized by the presence of a multimolecular film (eg, acacia, tragacanth, agar, chondrus) at the interface. There are two basic methods for preparing emulsions with the mortar and pestle. These are the *Wet Gum* (or so-called *English*) *Method* and the *Dry Gum* (or so-called *Continental*) *Method*.

The Wet Gum Method—In this method the emulsifying agent is placed in the mortar and dispersed in water to form a mucilage. The oil is added in small amounts with continuous trituration, each portion of the oil being emulsified before adding the next increment. Acacia is the most frequently used emulsifying agent when preparing emulsions with the mortar and pestle. When emulsifying a fixed oil, the optimum ratio of oil:water:acacia to prepare the initial emulsion is 4:2:1. Thus, the preparation of 60 mL of a 40% cod liver oil emulsion requires the following:

Cod liver oil	24 g
Acacia	6 g
Water, qs	60 mL

The acacia mucilage is formed by adding 12 mL of water to the 6 g of acacia in the mortar and triturating. The 24 g of oil is added in increments of 1–2 g and dispersed. The product at this stage is known as the *primary emulsion*, or *nucleus*. The primary emulsion should be triturated for at least 5 min, after which sufficient water is added to produce a final volume of 60 mL.

The Dry Gum Method—In this method, preferred by most pharmacists, the gum is added to the oil, rather than the water as with the wet gum method. Again, the approach is to prepare a primary emulsion from which the final product can be obtained by dilution with the continuous phase. If the emulsifier is acacia and a fixed oil is to be emulsified, the ratio of oil:water:gum is again 4:2:1.

Provided dispersion of the acacia in the oil is adequate, the dry gum method can almost be guaranteed to produce an acceptable emulsion. Because there is no incremental addition of one of the components, the preparation of an emulsion by this method is rapid.

With both methods the oil:water:gum ratio may vary, depending on the type of oil to be emulsified and the emulsifying agent used. The usual ratios for tragacanth and acacia are shown in Table XXIV.

The preparation of emulsions by both the wet and dry gum methods can be carried out in a bottle rather than a mortar and pestle.

Other Methods—An increasing number of emulsions are being formulated with synthetic emulsifying agents, especially of the nonionic type. The components in such a for-

Table XXIV—Usual Ratios of Oil, Water and Gum Used to Produce Emulsions

System	Acacia	Tragacanth
Fixed oils (excluding liquid petrolatum and linseed oil)	4	40
Water	2	20
Gum	1	1
Volatile oils, plus liquid petrolatum and linseed oil	2–3	20–30
Water	2	20
Gum	1	1

mulation are separated into those that are oil-soluble and those that are water-soluble. These are dissolved in their respective solvents by heating to about 70 to 75°. When solution is complete, the two phases are mixed and the product is stirred until cool. This method, which requires nothing more than two beakers, a thermometer and a source of heat, is necessarily used in the preparation of emulsions containing waxes and other high-melting-point materials that must be melted before they can be dispersed in the emulsion. The relatively simple methodology involved in the use of synthetic surfactant-type emulsifiers is one factor which has led to their widespread use in emulsion preparation. This, in turn, has led to a decline in the use of the natural emulsifying agents.

With hand homogenizers an initial rough emulsion is formed by trituration in a mortar or shaking in a bottle. The rough emulsion is then passed several times through the homogenizer. A reduction in particle size is achieved as the material is forced through a narrow aperture under pressure. A satisfactory product invariably results from the use of a hand homogenizer and overcomes any deficiencies in technique. Should the homogenizer fail to produce an adequate product, the formulation, rather than the technique, should be suspected.

For a discussion of the techniques and equipment used in the large-scale manufacture of emulsions, see Chapter 83.

Stability of Emulsions

There are several criteria which must be met in a well-formulated emulsion. Probably the most important and most readily apparent requirement is that the emulsion possess adequate physical stability; without this, any emulsion soon will revert back to two separate bulk phases. In addition, if the emulsified product is to have some antimicrobial activity (eg, a medicated lotion), care must be taken to ensure that the formulation possesses the required degree of activity. Frequently, a compound exhibits a lower antimicrobial activity in an emulsion than, say, in a solution. Generally, this is because of partitioning effects between the oil and water phases, which cause a lowering of the "effective" concentration of the active agent. Partitioning has also to be taken into account when considering preservatives to prevent microbiological spoilage of emulsions. Finally, the chemical stability of the various components of the emulsion should receive some attention, since such materials may be more prone to degradation in the emulsified state than when they exist as a bulk phase.

In the present discussion, detailed consideration will be limited to the question of physical stability. Reviews of this topic have been published by Garrett[63] and Kitchener and Mussellwhite.[64] For information on the effect that emulsification can have on the biologic activity and chemical stability of materials in emulsions, see Wedderburn,[65] Burt[66] and Swarbrick.[67]

The theories of emulsion stability have been discussed by Eccleston[68] in an attempt to understand the situation in both a simple O/W emulsion and complex commercial systems.

The three major phenomena associated with physical stability are

1. The upward or downward movement of dispersed droplets relative to the continuous phase, termed *creaming* or *sedimentation*, respectively
2. The aggregation and possible coalescence of the dispersed droplets to reform the separate, bulk phases.
3. Inversion, in which an O/W emulsion inverts to become a W/O emulsion, and *vice versa*.

Creaming and Sedimentation—Creaming is the upward movement of dispersed droplets relative to the continuous phase, while sedimentation, the reverse process, is the downward movement of particles. In any emulsion one process or the other takes place, depending on the densities of the disperse and continuous phases. This is undesirable in a pharmaceutical product where homogeneity is essential for the administration of the correct and uniform dose. Furthermore, creaming, or sedimentation, brings the particles closer together and may facilitate the more serious problem of coalescence.

The rate at which a spherical droplet or particle sediments in a liquid is governed by Stokes' law (Eq 35). While other equations have been developed for bulk systems, Stokes' equation is still useful since it points out the factors that influence the rate of sedimentation or creaming. These are the diameter of the suspended droplets, the viscosity of the suspending medium, and the difference in densities between the dispersed phase and the dispersion medium.

Usually, only the use of the first two factors is feasible in affecting creaming or sedimentation. Reduction of particle size contributes greatly toward overcoming or minimizing creaming, since the rate of movement is a square-root function of the particle diameter. There are, however, technical difficulties in reducing the diameter of droplets to below about 0.1 μm. The most frequently used approach is to raise the viscosity of the continuous phase, although this can be done only to the extent that the emulsion still can be removed readily from its container and spread or administered conveniently.

Aggregation and Coalescence—Even though creaming and sedimentation are undesirable, they do not necessarily result in the breakdown of the emulsion, since the dispersed droplets retain their individuality. Furthermore, the droplets can be redispersed with mild agitation. More serious to the stability of an emulsion are the processes of aggregation and coalescence. In aggregation (flocculation) the dispersed droplets come together but do not fuse. Coalescence, the complete fusion of droplets, leads to a decrease in the number of droplets and the ultimate separation of the two immiscible phases. Aggregation precedes coalescence in emulsions; however, coalescence does not necessarily follow from aggregation. Aggregation is, to some extent, reversible. While not as serious as coalescence, it will accelerate creaming or sedimentation, since the aggregate behaves as a single drop.

While aggregation is related to the electrical potential on the droplets, coalescence depends on the structural properties of the interfacial film. In an emulsion stabilized with surfactant-type emulsifiers forming monomolecular films, coalescence is opposed by the elasticity and cohesiveness of the films sandwiched between the two droplets. In spite of the fact that two droplets may be touching, they will not fuse until the interposed films thin out and eventually rupture. Multilayer and solid-particle films confer on the emulsion a high degree of resistance to coalescence, due to their mechanical strength.

Particle-size analysis can reveal the tendency of an emulsion to aggregate and coalesce long before any visible signs of instability are apparent. The methods available have been reviewed by Groves and Freshwater.[69]

Inversion—An emulsion is said to invert when it changes from an O/W to a W/O emulsion, or *vice versa*. Inversion sometimes can be brought about by the addition of an electrolyte or by changing the phase-volume ratio. For example, an O/W emulsion having sodium stearate as the emulsifier can be inverted by the addition of calcium chloride, because the calcium stearate formed is a lipophilic emulsifier and favors the formation of a W/O product.

Inversion often can be seen when an emulsion, prepared by heating and mixing the two phases, is being cooled. This takes place presumably because of the temperature-depen-

dent changes in the solubilities of the emulsifying agents. The phase inversion temperature, or PIT, of nonionic surfactants has been shown by Shinoda, *et al*[70] to be influenced by the HLB number of the surfactant. The higher the PIT value, the greater the resistance to inversion.

Apart from work on PIT values, little quantitative work has been carried out on the process of inversion; nevertheless, it would appear that the effect can be minimized by using the proper emulsifying agent in an adequate concentration. Wherever possible, the volume of the dispersed phase should not exceed 50% of the total volume of the emulsion.

Bioavailability from Coarse Dispersions

In recent years, considerable interest has focused on the ability of a dosage form to release drug following administration to the patient. Both the rate and extent of release are important. Ideally, the extent of release should approach 100%, while the rate of release should reflect the desired properties of the dosage form. For example, with products designed to have a rapid onset of activity, the release of drug should be immediate. With a long-acting product, the release should take place over several hours, or days, depending on the type of product used. The rate and extent of drug release should be reproducible from batch to batch of the product, and should not change during shelf life.

The principles on which biopharmaceutics is based are dealt with in some detail in Chapters 35 to 37. While most published work in this area has been concerned with the bioavailability of solid dosage forms administered by the oral route, the rate and extent of release from both suspensions and emulsions is important and so will be considered in some detail.

Bioavailability from Suspensions—Suspensions of a drug may be expected to demonstrate improved bioavailability compared to the same drug formulated as a tablet or capsule. This is because the suspension already contains discrete drug particles, whereas tablet dosage forms must invariably undergo disintegration in order to maximize the necessary dissolution process. Frequently, antacid suspensions are perceived as being more rapid in action and therefore more effective than an equivalent dose in the form of tablets. Bates, *et al*[71] observed that a suspension of salicylamide was more rapidly bioavailable, at least during the first hour following administration, than two different tablet forms of the drug; these workers were also able to demonstrate a correlation between the initial *in vitro* dissolution rates for the several dosage forms studied and the initial rates of *in vivo* absorption. A similar argument can be developed for hard gelatin capsules, where the shell must rupture or dissolve before drug particles are released and can begin the dissolution process. Such was observed by Antal, *et al*[72] in a study of the bioavailability of several doxycycline products, including a suspension and hard gelatin capsules. Sansom, *et al*[73] found mean plasma phenytoin levels higher after the administration of a suspension than when an equivalent dose was given as either tablets or capsules. It was suggested that this might have been due to the suspension having a smaller particle size.

In common with other products in which the drug is present in the form of solid particles, the rate of dissolution and thus potentially the bioavailability of the drug in a suspension can be affected by such factors as particle size and shape, surface characteristics, and polymorphism. Strum, *et al*[74] conducted a comparative bioavailability study involving two commercial brands of sulfamethiazole suspension (Product A and Product B). Following administration of the products to 12 normal subjects and taking blood samples at predetermined times over a period of 10 hr, the workers found no statistically significant difference in the extent of drug absorption from the two suspensions. The absorption rate, however, differed, and from *in vitro* studies it was concluded that product A dissolved faster than product B and that the former contained more particles of smaller size than the latter, differences that may be responsible for the more rapid dissolution of particles in product A. Product A also provided higher serum levels in *in vivo* tests half an hour after administration. The results showed that the rate of absorption of sulfamethiazole from a suspension depended on the rate of dissolution of the suspended particles, which in turn was related to particle size. Previous studies[75,76] have shown the need to determine the dissolution rate of suspensions in order to gain information as to the bioavailability of drugs from this type of dosage form.

The viscosity of the vehicle used to suspend the particles has been found to have an effect on the rate of absorption of nitrofurantoin but not the total bioavailability. Thus Soci and Parrott were able to maintain a clinically acceptable urinary nitrofurantoin concentration for an additional two hours by increasing the viscosity of the vehicle.[77]

Bioavailability from Emulsions—There are indications that improved bioavailability may result when a poorly absorbed drug is formulated as an orally administered emulsion. However, little study appears to have been made in direct comparison of emulsions and other dosage forms such as suspensions, tablets, and capsules; thus it is not possible to draw unequivocal conclusions as to advantages of emulsions. If a drug with low aqueous solubility can be formulated so as to be in solution in the oil phase of an emulsion, its bioavailability may be enhanced. It must be recognized, however, that the drug in such a system has several barriers to pass before it arrives at the mucosal surface of the gastrointestinal tract. For example, with an oil-in-water emulsion, the drug must diffuse through the oil globule and then pass across the oil/water interface. This may be a difficult process, depending on the characteristics of the interfacial film formed by the emulsifying agent. In spite of this potential drawback, Wagner, *et al*[78] found that indoxole, a nonsteroidal anti-inflammatory agent, was significantly more bioavailable in an oil-in-water emulsion than in either a suspension or a hard gelatin capsule. Bates and Sequeira[79] found significant increases in maximum plasma levels and total bioavailability of micronized griseofulvin when formulated in a corn oil/water emulsion. In this case, however, the enhanced effect was not due to emulsification of the drug in the oil phase *per se* but more probably because of the linoleic and oleic acids present having a specifical effect on gastrointestinal motility.

References

1. Semat H: *Fundamentals of Physics*, 3rd ed, Holt-Rinehart-Winston, New York, 1957.
2. Michaels AS: *J Phys Chem 65*. 1730, 1961.
3. Zisman WA: *Adv Chem Ser 43*· 1, 1964.
4. Titoff Z: *Z Phys Chem 74*. 641, 1910.
5. Lowell S: *Introduction to Power Surface Area*, Wiley-Interscience, New York, 1979.
6. Osipow LI: *Surface Chemistry Theory and Industrial Applications*, Reinhold, New York, 1962
7. Langmuir I: *J Am Chem Soc 39* 1848, 1917.
8. Giles CH: In EH Lucassen-Reynders, ed, *Anionic Surfactants*, Marcel Dekker, New York, 1981, Ch 4.
9. Weiser HB: *A Textbook of Colloid Chemistry*, Elsevier, New York, 1949.
10. Ter-Minassian-Saraga L: *Adv Chem Ser 43* 232, 1964.
11. Schott H, Martin AN. In Dittert LW, ed, *American Pharmacy*, 7th ed, Lippincott, Philadelphia, Chap 6, 1974.

12. Shinoda K, Nakagawa T, Tamamushi BI, Isemura T: *Colloidal Surfactants*, Academic, New York, Chap 2, 1963.
13. Sjoblom L. In Shinoda K, ed, *Solvent Properties of Surfactant Solutions*, Marcel Dekker, New York, Chap 5, 1967.
14. Prince, LM: *Microemulsions—Theory and Practice*, Academic, New York 1977.
15. Shinoda K, Friberg S: *Adv Colloid Interface Sci 4*: 281, 1975.
16. Friberg SE, Venable RV in Becher P, ed: *Encyclopedia of Emulsion Technology*, vol 1, Marcel Dekker, Chapt 4, New York, 1983.
17. Overbeek JThG: *Disc Faraday Soc 65*. 7, 1978.
18. Davis SS. In Bundgaard H, Hansen AB, Kofod H, eds: In *Optimization of Drug Delivery*, Alfred Benzon Symposium 17, Munksgaard, Copenhagen, 1982.
19. Fendler JH, Fendler EJ: *Catalysis in Micellar and Macromolecular Systems*, Academic, New York, 1975.
20. Mackay RA: *Adv Colloid Interface Sci 15*, 131, 1981.
21. Kruyt HR: *Colloid Science*, vols I and II, Elsevier, Houston, 1952 and 1949.
22. Alexander AE, Johnson P: *Colloid Science*, Oxford University Press, Oxford, 1949.
23. Ross S, Morrison ID: *Colloidal Systems and Interfaces*, Wiley, New York, 1988.
24. Mysels KJ: *Introduction to Colloid Chemistry*, Wiley-Interscience, New York, 1959
25. Shaw DJ: *Introduction to Colloid and Surface Chemistry*, 3rd ed, Butterworths, London, 1980.
26 Vold RD, Vold MJ: *Colloid and Interface Chemistry*, Addison-Wesley, Reading, MA., 1983.
27. Hiemenz PC. *Principles of Colloid and Surface Chemistry*, Marcel Dekker, New York, 1986.
28. von Weimarn PP. In Alexander J, ed: *Colloid Chemistry*, vol I, Chemical Catalog Co (Reinhold), New York, 1926. See also *Chem Rev 2* 217, 1926.
29. Lachman L *et al* *Theory and Practice of Industrial Pharmacy*, 3rd ed, Lea & Febiger, Philadelphia, 1986
30. Tubis M, Wolf W, eds: *Radiopharmacy*, Wiley-Interscience, New York, 1976.
31. LaMer VK, Dinegar RH. *J Am Chem Soc 72* 4847, 1950.
32 Matijevic E: *Acc Chem Res 14*. 22, 1981 and *Ann Rev Mater Sci 15* 483, 1985.
33 Florence AT, Attwood D: *Physicochemical Principles of Pharmacy*, Chapman & Hall, New York, 1982.
34. Gutch CF, Stoner MH: *Review of Hemodialysis*, Mosby, St. Louis, 1975.
35. Schott H, Martin AN. In Dittert LW, ed: *American Pharmacy*, 7th ed, Lippincott, Philadelphia, 1974.
36. Parks GA: *Chem Rev 65* 177, 1965.
37. Schott H: *J Pharm Sci 66*: 1548, 1977.
38. Sonntag H, Strenge K: *Coagulation and Stability of Disperse Systems*, Halstead, New York, 1972.
39. Hough DB, Thompson L. In Schick MJ, ed: *Nonionic Surfactants-Physical Chemistry*, 2nd ed, Marcel Dekker, Chapt 11, New York, 1987.
40. Vincent B: *Adv Colloid Interface Sci 4* 193, 1974.
41. Hunter RJ: *Zeta Potential in Colloid Science*, Academic, New York, 1981.
42. Davies JT, Rideal EK: *Interfacial Phenomena*, 2nd ed, Academic, New York, 1963
43. Bier M, ed: *Electrophoresis*, vols I and II, Academic, New York, 1959 and 1967.
44. Shaw DJ: *Electrophoresis*, Academic, New York, 1969.
45. Cawley LP. *Electrophoresis and Immunoelectrophoresis*, Little-Brown, Boston, 1969.
46. Catsimpoolas N, ed: *Isoelectric Focusing and Isotachophoresis*, *Ann NY Acad Sci 209*. June 15, 1973.
47. Morawetz H: *Macromolecules in Solution*, 2nd ed, Wiley-Interscience, New York, 1975.
48 Veis A: *The Macromolecular Chemistry of Gelatin*, Academic, New York, 1964.
49. Ward AG, Courts A, eds: *The Science and Technology of Gelatin*, Academic, Chapt 6, New York, 1977.
50. Haines BA, Martin A: *J Pharm Sci 50* 228, 753, 756, 1961.
51. Matthews BA, Rhodes CT: *J Pharm Pharmacol 20* Suppl: 204S, 1968.
52. Matthews BA, Rhodes CT: *J Pharm Sci 57* 569, 1968.
53. *Ibid 59* 521, 1970.
54. Schneider W, *et al*: *Am J Pharm Ed 42*: 280, 1978.
55. Tingstad J *et al*: *J Pharm Sci 62*: 1361, 1973.
56. Davies JT: *Proc Intern Congr Surface Activity*, 2nd, London, 426, 1957.
57. Sherman P: In *Emulsion Science*, Chap 4, Academic, New York, 1968.
58. Rogers JA: *Cosmet Toiletries 93*: 49, July, 1978.
59. Griffin WC: *J Soc Cos Chem 1*: 311, 1949.
60. Griffin WC: *J Soc Cos Chem 5*: 249, 1954.
61. Davies JT, Rideal EK: *Interfacial Phenomena*, Chap 8, Academic, New York, 1961. Davies JT: *Proc Intern Congr Surface Activity*, 2nd, London, 426, 1957.
62. Schott J: *J Pharm Sci 60*: 649, 1971.
63. Garrett ER: *J Pharm Sci 54* 1557, 1965.
64. Kitchener JA, Mussellwhite PR. In *Emulsion Science*, Academic, New York, Chap 2, 1968.
65. Wedderburn DL. In *Advances in Pharmaceutical Sciences*, vol 1, Academic, London, 195, 1964.
66. Burt BW: *J Soc Cosm Chem 16*: 465, 1965.
67. Swarbrick J: *Ibid 19*: 187, 1968.
68. Eccleston GM: *Cosmet Toiletries*: 101, 73, Nov 1986.
69. Groves MJ, Freshwater DC: *J Pharm Sci 57*: 1273, 1968.
70. Shinoda K, Kunieda H: In *Encyclopedia of Emulsion Technology*, Chap 5, Marcel Dekker, New York, 1983.
71. Bates TR *et al*. *J Pharm Sci 58* 1468, 1969
72. Antal EJ *et al*: *Ibid 64*. 2015, 1975.
73. Sansom LN *et al* *Med J Aust* 1975(2): 593.
74. Strum JD *et al*. *J Pharm Sci 67*: 1659, 1978.
75. Bates TR *et al*: *Ibid 62*. 2057, 1973.
76. Howard SA *et al* *Ibid 66*: 557, 1977.
77. Soci MM, Parrott EL: *Ibid 69*: 403, 1980.
78. Wagner JG *et al*: *Clin Pharmacol Ther 7*. 610, 1966.
79. Bates TR, Sequeira JA: *J Pharm Sci 64*: 793, 1975.

Bibliography

Interfacial Phenomena
Adamson AW: *Physical Chemistry of Surfaces*, 4th ed, Interscience, New York, 1982.
Davis JT, Rideal EK: *Interfacial Phenomena*, 2nd ed, Academic, New York, 1963.
Hiemenz, PC: *Principles of Colloid and Surface Chemistry*, 2nd ed, Marcel Dekker, New York, 1986.
Shaw DJ: *Introduction to Colloid and Surface Chemistry*, Butterworths, London, 1980.
Colloidal Dispersions
Particle Phenomena and Coarse Dispersions
Davies JT, Rideal EK: *Interfacial Phenomena*, Academic, New York, 1963.
Osipow LI: *Surface Chemistry*, Reinhold, New York, 1962.
Hiemenz PC: *Principles of Colloidal and Surface Chemistry*, Marcel Dekker, New York, 2nd ed, 1986.
Matijevic E, ed: *Surface and Colloid Science*, vols 1–4, Wiley, New York, 1971.
Cadle RD: *Particle Size*, Reinhold, New York, 1965.
Parfitt G: *Dispersion of Powders in Liquids*, Applied Science, 1973.
Adamson AW: *Physical Chemistry of Surfaces*, 4th ed, Wiley-Interscience, New York, 1980.
Fowkes FM, ed: *Hydrophobic Surfaces*, Academic, New York, 1969
Sherman P: *Rheology of Emulsions*, Macmillan, New York, 1963
Becher P: *Emulsions Theory and Practice*, 2nd ed, Reinhold, New York, 1965.
Vold RD, Vold MJ, *Colloid and Interface Chemistry*, Addison-Wesley, Reading, Mass, 1983.
Becher P: *Encyclopedia of Emulsion Technology*, Vols 1 to 3, Marcel Dekker, New York, 1983–1988.

CHAPTER 28

Clinical Analysis

Robert D Smyth, PhD
Vice President, Pharmaceutical Development

Lorraine Evans, BS, H(ASCP)
Clinical Pathology
Pharmaceutical Research & Development Div
Bristol-Myers Company
Syracuse, NY 13221

The characterization and quantitation of the various components of blood, urine and other body fluids are the primary functions of the clinical laboratory. The major divisions of clinical analysis are clinical biochemistry, hematology, blood-bank technology, histopathology, immunology and microbiology. The accurate diagnosis of disease and determination of a potential therapeutic regimen frequently are based on the laboratory analysis of blood, urine, feces, gastric secretions or cerebrospinal fluid. Modern medical practice is tending toward greater reliance on laboratory results as definitive measures of pathological or normal states.

The pharmacist should familiarize himself with the basic principles involved in sample collection, analysis and diagnostic significance of the various clinical parameters. His role in community health necessitates his comprehension of the methodology and diagnostic value of clinical laboratory procedures. The influence of various drugs and drug interactions on these parameters must be considered in both the clinical and drug-abuse situation.

Hematology

The determination of the morphological, physiological and biochemical properties of peripheral blood and the blood-forming organs (hematopoietic system) is a function of the hematology laboratory. The functional categories of hematology are (1) analysis of cellular elements, and specific biochemical and physiological parameters of peripheral blood and the hematopoietic system, (2) blood-coagulation analysis and (3) blood-bank technology.

Peripheral blood is a biphasic liquid tissue system of cellular elements suspended in a liquid plasma phase. The cellular phase comprises about 45% of the blood volume and contains erythrocytes (red blood cells, RBC), leukocytes (white blood cells, WBC) and thrombocytes (platelets). The plasma phase is primarily water (90 to 92%) and protein (7%).

The hematological analysis of blood is concerned primarily with enumeration and differentiation of the various cellular elements. An analysis of the hematopoietic system (eg, bone marrow and lymphoid tissue) determines the status of blood-cell precursors in these tissues. Determinations of specific biochemical (hemoglobin) and physiological (blood or plasma volume) parameters are performed in a complete evaluation of the erythron system (blood and marrow RBC and their precursors). The normal hematological values in the adult are presented in Table I.

The authors acknowledge the assistance of Dr Joseph P Uscavage of Rorer Group Ltd, in the preparation of the *Microbiology* section and Dr Alfred H Free of the Ames Co for the *Urinalysis* section.

Erythrocytes and Hemoglobin—The erythrocytic system is composed of the mature erythrocytes in peripheral blood and their precursors in bone marrow. The precursors of erythrocytes, as found in the erythropoietic system (red bone marrow), are classified as to the degree of nucleation and characteristics of cytoplasmic constituents. The sequence of erythrocyte formation in bone marrow—based on the gradual denucleation of the cell, generation of the chromatin structure and changes in nucleolar structure and cytoplasmic constituents—is as follows:

pronormoblast → basophilic normoblast → polychromatic normoblast → orthochromatic normoblast → polychromatophilic erythrocyte → erythrocyte.

The first four types are nucleated and normally are seen only in bone marrow. In normal erythrocyte formation these immature bone-marrow cells are designated as *normoblastic* or *normocytic*. In pernicious anemia and related conditions they become abnormally large and are designated *megaloblastic* or *megalocytic*. In iron-deficiency anemia, these cells become abnormally small and are designated *microblastic* or *microcytic*—of the iron-deficiency type.

Table I—Normal Hematological Values in Man[1]

	Normal Value	Normal Range of Values
Erythrocytes (cu mm × 10^6)		
Male	5.4	4.6–6.2
Female	4.8	4.2–5.6
Reticulocytes (cu mm × 10^3)	50	10–100
Hemoglobin (g%)		
Male	16.0	14.0–18.0
Female	14.0	12.0–16.0
Hematocrit (%)		
Male	47.0	40.0–54.0
Female	42.0	37.0–47.0
Mean corpuscular volume (μm^3)	87	82–92
Mean corpuscular hemoglobin (pg)	29	27–31
Mean corpuscular hemoglobin concentration (%)	34	32–36
Mean corpuscular diameter (μm)	7.3	6.7–7.7
Leukocytes (cu mm × 10^3)	7.0	5.0–10.0
Leukocyte differential (%)		
Neutrophils	63	57–67
Eosinophils	1	1–3
Basophils	1	0–1
Lymphocytes	30	25–33
Monocytes	5	3–7
Platelets (cu mm × 10^6)	3.0	1.4–6.0
Erthrocyte sedimentation rate (Wintrobe), (mm/hr)		
Male	4	0–9
Female	10	0–20

Normal blood contains 0.5 to 1.5% of circulating erythrocytes as reticulocytes. These cells contain a fine network of basophilic reticulum that is demonstrable on staining with a vital dye such as brilliant cresyl blue. The number of these cells in the blood is a measure of effective erythropoiesis. High circulating-reticulocyte values are an index of erythropoietic activity and are found in the first few days of life, after hemorrhage and after treatment of iron- or vitamin B_{12}-deficiency anemias.

The normal *erythrocyte* (normocyte) is a flexible, elastic, biconcave, enucleated structure with a mean diameter of 7.3 μm and a thickness near 2.2 μm. The chemical constituents of the red blood cell include water (63%), lipids (0.5%), glucose (0.8%), minerals (0.7%), nonhemoglobin protein (0.9%), methemoglobin (0.5%) and hemoglobin (33.6%). The primary function of the erythrocyte is transport of oxygen and carbon dioxide. The red cell membrane, a dynamic, semipermeable component of the cell, is associated with energy metabolism in the maintenance of the permeability characteristics of the cell to various cations (Na^+, K^+) and anions (Cl^-, HCO_3^-). The stroma of insoluble material which remains after red-cell disruption (hemolysis) constitutes 2 to 5% of the wet-cell weight; it is primarily protein (40 to 60%) and lipid (10 to 12%). The membrane includes stromatin (a fibrous or structural protein) and mucopolysaccharides associated with A, B and O blood-group substances. The lipid fractions include phosphatides (lecithin, cephalin), cholesterol, cholesterol esters, neutral fats, cerebrosides and sialic acid glycoproteins.

Erythrocytes may be enumerated by either visual or electronic procedures. In the visual procedures, a measured quantity of blood is diluted with a fluid which is isotonic with blood and will prevent its coagulation. The diluted blood is then placed in a counting chamber (hemocytometer), and the number of cells in a circumscribed area is enumerated microscopically. Hayem's solution (sodium sulfate, 2.5 g; sodium chloride, 0.25 g; mercuric chloride, 0.25 g; distilled water, 100 mL), Toison's fluid (sodium sulfate, 8 g; sodium chloride, 1 g; methyl violet, 0.025 g; glycerin, 30 mL; distilled water, 180 mL) or 0.9% sodium chloride are used as diluting fluids. The overall error of this method is about 8%.

A greater degree of accuracy and reproducibility can be achieved by erythrocyte enumeration in an electronic counting apparatus; eg, Coulter Counter or Ortho cell counters. The Coulter method (Fig 28-1) determines the number and size of particles suspended in an electrically conductive liquid. The blood cells traverse a small aperture and displace their own volume in the diluent as to produce a change in resistance between the electrodes; the magnitude of the voltage pulse is proportional to cell volume, and the resultant pulses are then amplified, scaled and automatically counted.

In the Ortho ELT-8 technique (Fig 28-2), the principles of laser flow cytometry are used to count cells. Hydrodynamic focusing and laminar flow are combined in the system to count a large number of individual cells. Light focused by a helium-neon laser is scattered by the cells as they pass through the flow channel. The scattered light is monitored by a photoelectric sensor and transfers the electrical pulses which are processed by the systems circuitry. In addition to increased counting speed, the overall error of the electronic procedures is reduced to about 1%.

The *hematocrit value* is also a measure of the erythrocyte portion of blood. A sample of blood containing an anticoagulant is placed in a graduated hematocrit capillary tube, centrifuged and the volume ratio of packed red cells to total blood volume (hematocrit value) determined. The centrifuged sample appears as a red layer of packed erythrocytes over which is found an off-white layer of packed leukocytes and platelets, and a supernatant plasma phase. The hematocrit value is an index of both the number and size of the red cells.

Hemoglobin, a conjugated hemoprotein with an approximate molecular weight of 67,000, contains basic proteins, the globins and ferroprotoporphyrin (heme). It is essentially a tetramer, consisting of four peptide chains, to each of which is bound a heme group. Heme, which constitutes about 4% of the weight of the molecule, consists of a divalent iron atom in the center of a pyrrole-porphyrin structure. Four distinct polypeptide chains (α, β, γ, δ) can be incorporated into hemoglobin. Normal adult hemoglobin is HbA = $\alpha_2^A\beta_2^A$. Fetal hemoglobin contains 2α and 2γ chains and is designated HbF = $\alpha_2^A\gamma_2^F$.

Differences in the structural sequences of amino acids in the peptide portion of the hemoglobin molecules are controlled genetically and are responsible for different types of hemoglobin. Based on the characteristic mobility of the hemoglobin, in an electric field (electrophoresis) on starch, paper, cellulose acetate, agar or acrylamide gel media, many hemoglobin types have been recognized (see Chapter 29). Only types P, F and A_1-A_4 are considered normal. Sickle-cell anemia and β-thalassemia are hemolytic anemias associated with abnormal hemoglobins (ie, Type S in sickle-cell anemia and abnormal production of the β chain in β-thalas-

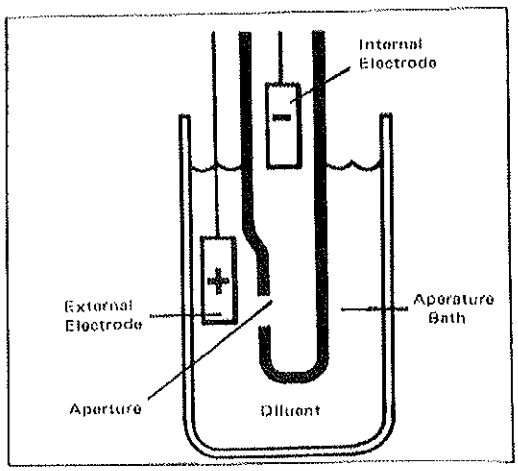


Fig 28-1. Coulter-counting cells by electronic impedance (courtesy Coulter Electronics).

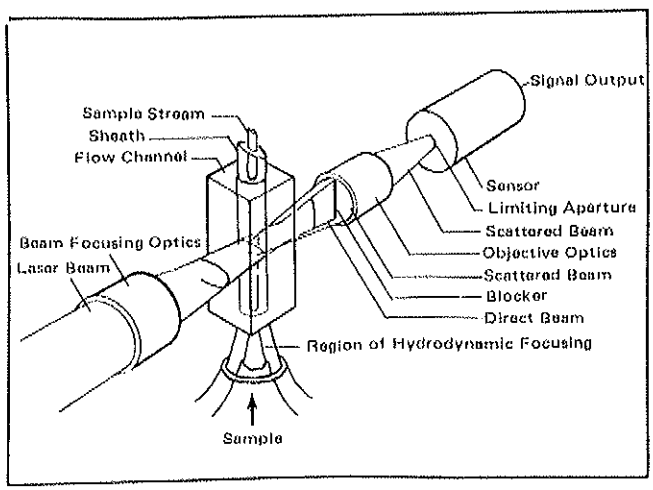


Fig 28-2. Ortho ELT-8-Method of scattered light detection and hydrodynamic focusing for cell counting (courtesy, Clinical Instrument Systems, Oct 1980).

semia). In homozygous *HbS disease* sickling of the red cells is due to the low solubility of the abnormal hemoglobin in its reduced state, with the production of semicrystalline bodies (tactoids), which distort and elongate the cells. In the sickle-cell trait (heterozygous), the blood smear shows no sickle cells. In the homozygous condition, HbS accounts for nearly all of the hemoglobin with small amounts of HbF. In the heterozygous condition, HbS constitutes 50% or less of the hemoglobin, with the balance as HbA.

The detection of sickle-cell disease is performed by microscopic observation of the induction of red-cell sickling in the presence of a reducing agent such as sodium metabisulfite or by quantitative determination of urea-dispersible turbidity induced by dithionite following reduction of HbS to deoxy-HbS in RBC lysates. The microscopic procedure will detect only homozygotes, whereas HbAS and HbS and its structural variant HbC-Harlem both are detected in the urea-dithionite technique. Commercial qualitative test kits are available for detecting sickle-cell trait and anemia by solubility determinations. All hemoglobins positive to the dithionite test must be electrophorized (cellulose acetate, citrate agar or starch gel) to differentiate HbS from HbC and thalassemia traits. Drugs causing hemolysis in glucose 6-phosphate dehydrogenase (G6PD) deficiency include sulfones, nitrofurans, chloroquine, dimercaprol, nalidixic acid and probenecid.

The *hemoglobin* concentration is measured spectrophotometrically after lysis of whole blood and conversion of hemoglobin to hematin, oxyhemoglobin or cyanmethemoglobin. The addition of a strong base (NaOH) to pH 10 converts oxyhemoglobin, carboxyhemoglobin and methemoglobin to hematin, which can be estimated photometrically. Weaker bases (Na_2CO_3 or NH_4OH) convert hemoglobin to oxyhemoglobin for analysis.

Total hemoglobin is measured also by conversion to cyanmethemoglobin using alkaline sodium cyanide-potassium ferricyanide reagent. Hemoglobin standards certified by the Clinical Standards Committee of the College of American Pathologists are used in these procedures, and all results are expressed as "g hemoglobin per 100 mL blood."

In the normal state, the oxygen consumption of the RBC is low and it is involved in the conversion of hemoglobin to oxidized (Fe^{+4}) methemoglobin (HbM) which cannot bind oxygen. The normal balance of HbM (<0.5%) is maintained by two enzyme systems—NADH and NADPH methemoglobin reductases. An inherited deficiency of the RBC enzyme, G6PD. This will decrease the rate of reduction of glutathione and methemoglobin, make the cell more vulnerable to oxidative attack and result in susceptibility to drug-induced or immune-mediated nonspherocytic hemolytic anemia. G6PD deficiency is found predominantly in Mediterranean peoples, Southeast Asians, Africans and American negroes. The enzyme can be quantitated spectrometrically or by fluoronephelometry by measuring the rate of reduction of nicotinamide adenine dinucleotide phosphate (NADP) in the presence of G6PD. Presumptive screening tests based on reduced glutathione (GSH) content of blood before and after incubation with acetylphenylhydrazine also are used.

Erythrocyte count, hemoglobin content and hematocrit value are used to determine various blood indices in the diagnosis and treatment of anemia. These measurements are:

Mean corpuscular volume {MCV (μm^3)} =

$$\frac{\text{Hematocrit (\%)} \times 10}{\text{Erythrocyte count (millions/cu mm)}}$$

Mean corpuscular hemoglobin {MCH (pg)} =

$$\frac{\text{Hemoglobin (g/100 mL)} \times 10}{\text{Erythrocyte count (millions/cu mm)}}$$

Mean corpuscular hemoglobin concentration {MCHC(%)} =

$$\frac{\text{Hemoglobin (g/100 mL)} \times 100}{\text{Hematocrit (\%)}}$$

An additional parameter used to characterize red-cell variation is the red-cell distribution width (RDW) determined on the Coulter S-Plus II. The RDW is calculated directly by the standard deviation and coefficient of variation from a red-cell histogram on the S-Plus II. The difference in cell size may be used to monitor patients with pernicious or hemorrhagic anemia.

Anemias are classified as to red-cell volume and hemoglobin concentration. *Macrocytic* (large cell: MCV > 94), *normocytic* (normal cell: MCV, 82 to 92), or *microcytic* (small cell: MCV < 80) are the classifications according to cell volume. Cellular hemoglobin concentration categorizes the cells as to *hyperchromic* (MCHC > 38), *normochromic* (MCHC = 32 to 36), or *hypochromic* (MCHC < 30). Examples of anemias:

I. Hypochromic Microcytic—erythroid normoblastic anemia in bone marrow
   - A. Iron Deficiency—low hemoglobin (Hbg) and RBC, low serum iron, high total iron binding capacity, absent hemosiderin.
     1. Dietary—low iron intake
     2. Intestinal problems—decreased iron absorption
     3. Pregnancy, infants—increased iron requirements
     4. Iron loss—due to chronic hemorrhage, parasitic infections, GI tract lesions, excess menstrual bleeding.
   - B. Hereditary Sideroblastic—defect in the heme synthesis, an inability to utilize ingested iron.
   - C. Thalassemia—genetic abnormality which produces normal to increased $HbgF$ and/or $HbgA_2$.

II. Normochromic Normocytic
   - A. Hemolytic—increased destruction of erythrocytes.
     1. Autoimmune hemolytic
     2. Cold agglutinin hemolytic
     3. Mechanical destruction of RBCs
     4. Paroxysmal Nocturnal hemoglobinuria
     5. Lymphomas and Hodgkin's disease
     6. Infections
   - B. Hemoglobinopathies—abnormalities in structure of alpha or beta chains of hemoglobin molecule; normoblastic erythroidhyperplasia in bone marrow.
     1. Sickle-cell
     2. Hemolysis
     3. Hemoglobin CC
   - C. Acute Hemorrhage
   - D. Other
     1. Aplastic Anemia, Leukemia, Malignancy
     2. Renal failure and drug-related anemias caused by chloramphenicol and antineoplastic drugs.

III. Normochromic Macrocytic—due to deficiency of vitamin B_{12} or folate; bone marrow is hypercellular with increased erythroid precursors.
   1. Pernicious
   2. Sideroblastic
   3. Sprue—total iron-binding capacity is decreased; hemosiderin is increased in the bone marrow.
   4. Pregnancy

Determinations of the suspension stability of whole blood and erythrocyte fragility are useful adjuncts in the diagnosis of various diseases.

The *erythrocyte sedimentation rate* (ESR) is an estimate of the suspension stability of red blood cells in plasma; it is related to the number and size of the red cells and to the relative concentration of plasma proteins, especially fibrinogen and the α- and β-globulins. This test is performed by determining the rate of sedimentation of blood cells in a standard tube. Normal blood ESR is 0 to 15 mm/hour. Increases are an indication of active but obscure disease processes such as tuberculosis and ankylosing spondylitis. ESR is affected by anemia and does not respond linearly with changes in asymmetrical macromolecules such as fibrinogen and globins.

The *zeta sedimentation ratio* (ZSR) technique overcomes these disadvantages. It is based on a measure of the closeness with which RBC will

approach each other after standardized cycles of dispersion and compaction.

The *erythrocyte fragility test* is based on resistance of cells to hemolysis in decreasing concentrations of hypotonic saline.

Increased osmotic fragility of the red cells is associated with various types of spherocytosis and acquired hemolytic anemia; increased resistance has been observed in thalassemia, sickle-cell anemia and hypochromic anemia. The test can be performed manually by colorimetric estimation of hemoglobin released by hypotonic cell rupture or automatically in an instrument which continually records the increase in light transmittance through a suspension of red cells in a continuously decreasing salt gradient during dialysis.

Leukocytes—Mature *leukocytes* (white blood cells, WBC) in peripheral blood and their precursors in bone and lymphoid tissue comprise the leukocytic system. Various types of leukocytes are found in normal blood. Differentiation of the lymphocytic, monocytic and granulocytic leukocyte types is based on cell size, color, chromatin structure and cytoplasm constituents.

The primary function of leukocytes is the development of the various defensive and reparative processes in inflammatory and immune-response mechanisms. The migration of leukocytes to the site of inflammation is associated with the release or activation of various biochemical substances (5-hydroxytryptamine, histamine, complement, immunoglobulins, prostaglandins, lysosomal enzymes). The tissue histiocyte or monocyte (macrophage) also can engulf and destroy foreign particles by the process of endocytosis and certain leukocyte types by phagocytosis.

The chemical composition of the leukocyte includes water (82%), nucleoprotein, phospholipids and trace minerals. Enzyme content, glycogen and histamine levels vary in the different types of white cells. Deficiency in enzymes associated with glycolytic metabolism (hexokinase) and increases in phosphomonoester hydrolases (alkaline phosphatase) have been observed in leukocytes of certain leukemia patients.

The precursors of granulocytic leukocytes are found in bone marrow and are classified according to the degree of cytoplasmic granulation, dye-affinity of the granules and shape of the nucleus (Schilling, Arneth or Cooke-Ponder Classification). As undifferentiated cells (myeloblasts) mature

promyelocyte → myelocyte → metamyelocyte → band leukocyte → segmented leukocyte

metachromatic granules appear in the cytoplasm (granulocytes). All segmented leukocytes are motile, a requirement for participation in the inflammatory or phagocytic processes.

In the mature *basophilic* and *eosinophilic leukocytes*, these granules develop an affinity for a basic or acidic dye, respectively; those cells containing granules which do not stain are called *neutrophils*. In peripheral blood, the mature granulocytic cells are designated *polymorphonuclear leukocytes—neutrophilic, eosinophilic or basophilic*.

The other types of white cells normally observed in peripheral blood have no granules and are classified as to size and shape into the *monocyte* and *lymphocyte*, which are formed in lymphoid tissue. The small lymphocyte is thymic-derived and is found in the circulation and germinal centers of lymphoid tissue. The origin of the large lymphocyte is a gut-associated lymphoid stem cell which can further differentiate into the immunoglobulin-producing plasmacyte. The interaction of thymic (T) and bone-marrow (B) lymphocytes is the basis for the development and maintenance of humoral and cellular immune mechanisms.

Leukocytes are enumerated by procedures similar to those used for erythrocytes. In the visual procedures the blood is diluted with a fluid (3% v/v acetic acid) which lyses the red cells, and the total leukocyte count is determined microscopically. Eosinophils also may be analyzed differentially with a diluting fluid which renders the red cells nonrefractile and invisible, and lyses the base-labile leukocytes, leaving the base-stable eosinophils intact. A suitable diluting fluid for this purpose is Pilot's Fluid (propylene glycol, 50 mL; distilled water, 40 mL; 1% phloxine, 10 mL; 10% sodium carbonate, 1 mL and heparin sodium, 100 units). Electronic-counting procedures are similar to those used for erythrocytes with the added advantages of speed, accuracy and reproducibility.

The normal adult leukocyte value is 5000 to 10,000 cells/cu mm. Values greater than 10,000 (*leukocytosis*) are encountered in the newborn infant, young children, after violent exercise, convulsive seizures of epilepsy, leukemia and cancer. Values of less than 5000 (*leukopenia*) are observed in certain microbial infections (eg, typhoid fever, measles, malaria, overwhelming septicemia), cirrhosis of the liver, pernicious anemia, radiation injury and replacement of marrow by malignant tissue.

A *differential count of the leukocytes* provides information as to the relative numbers of each type. A thin film of blood is prepared on a microscope slide stained with a polychromatic preparation such as the Leishman, Wright or Giemsa stain, and analyzed microscopically. Wright's stain contains polychromed methylene blue and eosin dyes; the erythrocytes are stained pink; the nuclei of the leukocytes, purplish-blue; neutrophilic granules, violet-pink; eosinophilic granules, red; basophilic granules, blue; and platelets, blue.

The recent introduction of automated systems for differential white-cell counts significantly reduce the errors inherent with the subjective nature of the visual counting procedure. Differentiation of the various cell types can be made on the basis of cytochemistry and staining properties of enzymes specific for a single cell type. The granules of neutrophils and eosinophils are stained by action of their peroxidases on 4-chloro-1-naphthol to form a colored quinone in the presence of a peroxide and further differentiated by the optimum pH for peroxidase activity between these two cell types. The monocytic lipase is used as a specific marker by the reaction of basic fuchsin with α-naphthol liberated by lipase on α-naphthylbutyrate substrate. The lymphocytes are not stained in this procedure but are measured by electronic sizing.

Automated differential WBC counts also have been obtained in systems which count large populations of cells by simultaneous measurement of two optical properties (axial light loss and/or narrow-angle scatter and/or multiple-wavelength fluorescence). Laser light also is used to differentiate cell size, granularity and volume of cells. The collected light measured by forward versus right-angle scatter is converted to a histogram giving the percent of lymphocytes, monocytes and granulocytes. Another system involves computer processing of two-dimensional images of the various cell types using an automatic scanning microscope.

Polymorphonuclear neutrophilic leukocytes (neutrophils, "polys") normally comprise 62% (50 to 67%) of the total leukocyte count. These cells are irregular in shape (10 to 15 μm in diameter) and usually contain a multilobated nucleus with fine, lightly stained cytoplasmic granules. An immature or juvenile form of neutrophil, with a band-shaped nonsegmented nucleus constitutes 3 to 5% of peripheral blood leukocytes. Increases in the relative percentage of these cells (neutrophilia) is observed in acute microbial infections (eg, meningitis, smallpox, poliomyelitis), metabolic disorders (diabetic acidosis, gout), drug intoxication (digitalis, epinephrine), vaccination, coronary thrombosis and malignant neoplasms.[2]

Polymorphonuclear eosinophilic leukocytes (eosinophils) normally comprise about 1 to 3% of total circulating white-blood cells. In appearance they are similar to the neutrophil with the exception of large, red-stained cytoplasmic granules. Eosinophilia has been observed in certain

skin diseases (psoriasis, eczema), parasitic infestations (pork round worm—trichinosis), certain hypersensitivity reactions, scarlet fever and pernicious anemia. Charcot-Leyden crystals, which are found in bronchial secretions from asthmatics, are derived from nucleoprotein-disintegration products of eosinophils.

Polymorphonuclear basophilic leukocytes (basophils) possess large cytoplasmic granules which stain a deep blue. These cells, which are primarily sources of blood heparin and histamine, constitute less than 1.0% of the leukocytes. Basophilic leukocytosis is seen in chronic myelocytic leukemia, hemolytic anemia and Hodgkin's disease. Basophilic leukopenia occurs following radiation or therapy with glucocorticoids.

Lymphocytes have a cell diameter from 7 to 10 μm (small) to 10 to 18 μm (large). They have a round, or slightly indented, deeply stained nucleus and normally comprise 25 to 33% of the leukocytes. Lymphocytosis is seen in infectious mononucleosis, lymphocytic leukemia, rickets and in most conditions associated with neutrophilic leukopenia (neutropenia).

Monocytes constitute 3 to 7% of the leukocytes. They are larger (12 to 20 μm) than the other leukocytes and possess an abundant, pale, bluish-violet-stained cytoplasm with a fine, reticulated chromatin structure in the nucleus. The monocytes (macrophages) phagocytize bacteria, parasitic protozoa, foreign particles and even erythrocytes. Monocytosis is seen in certain microbial infections (tuberculosis, typhus, malaria), Hodgkin's disease and monocytic leukemia.

Drug therapy frequently causes neutrophil dysfunction which can be characterized by a decreased number of mature neutrophils or a defect in cellular function resulting in the inability of the body to defend itself against infection. Drugs such as nitrogen mustard and chloramphenicol degenerate bone-marrow stem cells, and DNA synthesis is impaired by antimetabolites such as methotrexate and fluorouracil. Depolymerization of DNA is caused by procarbazine and alkylating agents. Mitosis is inhibited by colchicine and vinca alkaloids. The following outline lists drugs which cause granulocytopenia.[2]

- Nonchemotherapeutic
  - rifampin
  - fistocetin
  - benzene
  - nitrous oxide
  - ethanol
- Antithyroid
  - carbimazole
  - methimazole
  - thiouracil
- Diuretics
  - acetazolamide
  - chlorthalidone
  - chlorothiazide
  - ethacrynic acid
  - hydrochlorothiazide
  - mercurials
- Antihistamines
  - ethylenediamine
  - thenalidine
  - metaphenylene
  - pyribenzamine
- Phenothiazines
  - chlorpromazine
  - mepazine
  - methotrimeprazine
  - prochlorperazine
  - thoridizine
- Antibiotics
  - chloramphenicol
  - carbenicillin
  - griseofulvin
  - isoniazid
  - novobiocin
- Cardiovascular
  - diazoxide
  - procainamide
  - methyl dopa
  - quinidine
  - propranolol

As qualitative and quantitative changes in leukocytes in peripheral blood and their precursors in bone marrow and lymphatic tissue are associated with the various types of *leukemia*, this disease has been classified on the basis of the predominating type of leukocyte, ie, myelocytic (granulocytic), lymphocytic, monocytic or plasmacytic. Leukemia may be either acute or chronic and involve the replacement of bone-marrow elements by malignant cells, infiltration of the reticuloendothelial system, anemia, thrombocytopenia and hemorrhage. Leukemia usually is associated with an elevated WBC count and increase in the specific cell and its precursors in peripheral blood, but in certain instances there is an aleukemic blood picture with no evidence of leukocytosis. Leukocytes in acute leukemia are more immature ("blast"-type cells) than those encountered in the chronic type.

In many diseases of the hematopoietic system, it is necessary to examine the bone marrow to determine the rates of formation, maturation and release of blood cells into the peripheral circulation. Using a puncture biopsy needle, samples of *bone marrow* may be obtained from the sternum, iliac crest or proximal end of the tibia. Smears of marrow then are prepared, stained (Wright's stain or specialized histopathological procedure) and examined microscopically. The ratio of myeloid leukocyte to nucleated red cells in bone marrow, the presence of abnormal (*nonmyeloid*) cells, the number of platelet precursors (*megakaryocytes*), the signs of cell-maturation arrest and the presence of focal lesions are important factors in the diagnosis of various disease states.

Systemic lupus erythematosus (SLE) is a disease characterized by numerous clinical and pathological manifestations associated with various organs. Although the disease chiefly affects the lymphatic system, the cardiac, renal and articular systems also are involved. The diagnosis of this disease is based on the presence of an SLE-cell factor in the gamma-globulin fraction of blood in the diseased state. This factor dissolves the nuclei of leukocytes by depolymerization of deoxyribonucleic acid to form the SLE-body. If serum from patients with SLE is incubated with white cells, the "polys" will engulf the liberated SLE-body and form the typical SLE-cell with a characteristic progressive loss of nuclear detail. Drugs which cause SLE and produce a positive SLE-prep include hydralazine, procainamide, isoniazid and phenytoin.

These antibodies to nucleoprotein also can be detected by immunological techniques. In the double-antibody technique, the test serum containing antibodies to nuclear protein is incubated with a rat kidney slice (antigen). The second antibody is a fluorescein-labeled goat antihuman immunoglobulin (IgG) which combines with the human IgG bound to the antigen site in a positive test. The fluorescence is estimated by immunomicroscopy. Normal light-microscopy can be used if the goat-antihuman IgG is labeled with peroxidase.

Thrombocytes—The primary functions of *thrombocytes* (blood platelets) are the maintenance of hemostasis (arrest of blood flow from a vessel) and blood coagulation (clot formation). Platelets are oval to spherical in shape and have a mean diameter of 2 to 4 μm. They originate from an immature cell (megakaryocyte) in bone marrow and ranges of 140,000 to 450,000/cu mm have been reported in normal blood.

Adhesiveness, aggregation and agglutination are the principal physical properties of platelets responsible for hemostasis and coagulation reactions. Chemically, they contain protein (60%), lipid (15%) and carbohydrate (8.5%). Their content of serotonin, epinephrine and norepinephrine aids in promoting constriction at the site of injury. The release of "platelet thromboplastin," a cephalin-type phosphatide, and ADP are important in blood coagulation.

As of the present time, there is no satisfactory manual method for accurate enumeration of blood platelets. The size and physical properties of the platelet seriously deter the development of accurate and reproducible methodology. Indirect methods of analysis are based on the proportion of platelets to erythrocytes in a stained blood smear. Blood samples obtained directly from the fingertip puncture are diluted with an anticoagulant fluid which simultaneously will stain the platelets. The ratio of platelets to red cells then is determined microscopically and the number calculated from the predetermined red-cell count (normal 3 to 8 platelets/100 RBC). In the direct procedures, a sample of blood is obtained by venipuncture, placed in a siliconized

tube, diluted and subsequently analyzed by counting the platelets in a microscopic counting chamber using conventional or phase-microscopy apparatus. Suitable diluting fluids are the Rees-Ecker Fluid (sodium citrate, 3.8 g; formaldehyde, 0.22 mL; brilliant cresyl blue, 0.05 g; water, qs 100 mL) or Brecker Fluid (1% ammonium oxalate). Automated procedures for platelet counting have increased the accuracy to ±5 to 10%. Blood is collected in a special anticoagulant, diluted and centrifuged at specified speeds to obtain a "platelet-rich" supernatant fluid, which then is counted in an automated counting apparatus similar to those used for RBC counting.

Methods for counting platelets in whole blood include electronic impedence instruments and laser-optical counters using hydrodynamic focusing.[3] These new hematology multiparameter analyzers provide greater accuracy, precision and increased rate of analysis performed on a small volume of blood. The automated instruments provide precise platelet measurements for monitoring chemotherapy-induced thrombocytopenia and transfusion therapy.

Persistent increases in platelet count (*thrombocythemia* or *piastrenemia*) have been observed in chronic myelocytic leukemia, polycythemia, megakaryocytic hyperplasia and splenic atrophy. Acute or temporary increases in platelet values (*thrombocytosis*) are seen in trauma and asphyxiation.

Thrombocytopenia or a decrease in platelets to values less than 60,000/cu mm occurs in various purpuras or hemorrhagic states (idiopathic or symptomatic thrombocytopenic purpura). Inherited platelet defects include Glanzmann's thrombasthenia which is characterized by prolonged bleeding time and poor clot retraction, while Bernard-Soulier Syndrome and Von Willebrand's disease demonstrates defective platelet adhesiveness. Defects in the release reaction includes "Storage Pool Deficiency" and "Aspirin-like" syndrome.

A rare, inherited, structural and functional platelet abnormality is the *grey-platelet syndrome* characterized by large platelets lacking alpha granules and appearing grey on Wright's-stained peripheral blood smears. Patients have a history of bleeding, petechiae, easy bruising and expistaxis. Diagnosis is confirmed by radioimmunoassay procedures to detect levels of platelet-specific alpha-granule proteins.

Leukemia, extensive burns, splenic disorders and agents such as quinidine, sulfonamides, hydrochlorothiazide, diuretics, antiepileptics and neuropharmacological agents have been implicated in the etiology of symptomatic thrombocytopenia. Decreases in platelet count also are accompanied by morphological changes in the size, shape and cytoplasmic granulation of these cells and changes in adhesiveness and normal function in hemostasis and coagulation.

Studies on *platelet aggregation* have been of significant value in the study of platelet abnormalities and their role in disease states. The rate and extent of the aggregation and clotting response to adrenaline, ADP, collagen and thrombin have been measured by observing changes in optical density of platelet-rich plasma on adding of these agents or other test substances. Low amounts of ADP give reversible aggregation, while a biphasic-aggregation pattern occurs with intermediate concentrations of ADP or with epinephrine. The second phase is the release of the platelets' endogenous ADP. High concentrations of ADP result in an irreversible aggregation. Aspirin acts as an inhibitor of the intrinsic-platelet ADP and the collagen reaction.

Reticulocytes—In normal peripheral blood 0.5 to 1.5% of the erythrocytes possess a fine reticulum in the cytoplasm. In blood smears prepared with Wright's, Giemsa and other Romanowsky methods, basophilic stippling of the erythrocytes occurs in lead poisoning (*plumbism*). This is not to be confused with the basophilic staining of the reticulocyte which only can be seen when cells are stained by supravital procedures (mixture of dyes with wet blood prior to preparing of an air-dried blood smear). The observed granular filaments or reticulum of this immature erythrocyte are a result of endoplasmic coagulation by lipophilic dyes used in the supravital procedures. *Reticulocytes* are enumerated by supravital staining of fresh blood with an anticoagulant-dye solution.

The usual method of expression is

$$\% \text{ Retics} = \frac{\text{No of reticulocytes/1000 RBC}}{10} \quad (1)$$

The "corrected" reticulocyte count is calculated for a more meaningful clinical approach in the degree of anemia by expressing the percentage of reticulocytes per mm^3 of whole blood.

$$\frac{\text{Corrected reticulocyte count}}{} = \text{Reticulocyte count} \times \frac{\text{(Patient's hematocrit)}}{\text{(Normal hematocrit)}}$$

In indirect counting methods a thin film of the blood-dye mixture is prepared on a microscope slide, counterstained with Wright's stain and the reticulocytes enumerated in proportion to a predetermined erythrocyte count. In direct procedures, reticulocytes are enumerated in wet films without counterstaining. Suitable dyes are brilliant cresyl blue, methylene blue and Janus green. These methods are subject to a high counting error.

An increase in the number of reticulocytes is an index of accelerated hematopoiesis and is observed in acute hemorrhage or adequate therapeutic management of iron-deficiency or pernicious anemia. In cases of chronic blood loss or bone-marrow depression a decrease in reticulocytes is seen.

Blood-Volume and Erythropoietic Mechanisms—The mean red-cell mass in normal males is 2095 ± 384 mL (30 mL/kg), the average plasma volume is 2766 ± 459 mL (40 mL/kg) and the total blood volume is 4861 ± 795 mL (70 mL/kg). The specific determination of *red-cell mass* is estimated accurately by tagging erythrocytes with ^{51}Cr *in vitro* or ^{59}Fe *in vivo*. These isotopes are incorporated into the β-polypeptide (Cr) or porphyrin (Fe) of hemoglobin in the RBC and subsequent isotope dilution in blood after injection of tagged erythrocytes is used for calculation of red-cell mass. In hemolytic anemia there is also a decrease in the normal life span (108 to 120 days) of the erythrocyte as indicated by a decreased survival time of ^{51}Cr-tagged red cells in blood (refer to Chapter 33).

Plasma volume is estimated by measurement of hemodilution of IV-injected ^{125}I or ^{131}I human serum albumin. The activity of labeled albumin steadily decreases after injection due to the loss of albumin to the extravascular space. Estimates of zero-time radioactivity levels can be made by extrapolation of a typical first-order blood-level decay curve. Dyes (Evans Blue) and other isotopes are less satisfactory for accurate assessment of plasma volume. The total blood volume is equal to the red-cell mass and plasma volume.

Chronic expansion of the red-cell mass is seen in primary and secondary polycythemia associated with erythrocytosis due to hypoxia, tumors and renal disease. In these conditions, there is an increased hemoglobin and hematocrit and absolute increase in red-cell mass. In relative polycythemia the high hematocrit is due to contraction of the plasma volume. *Chronic expansion of the blood volume*, with a resultant decrease in hematocrit value and, in some cases, a "hemodilution" anemia, is seen in cardiac failure, normal pregnancy, hepatic cirrhosis, splenomegaly and arteriovenous fistula.

The metabolic defect in *pernicious anemia*, characterized by inadequate gastrointestinal absorption of vitamin B_{12}, is diagnosed readily by monitoring urinary radioactivity fol-

lowing oral administration of cyanocobalamin-^{57}Co with and without intrinsic factor. The percent recovery of the isotope in normal patients is 3 to 25% and in pernicious anemia 0 to 2.5%.

^{51}Cr-tagged erythrocytes also are used in studying the effects of various compounds, such as the nonsteroidal anti-inflammatory drugs, on *gastrointestinal (GI) bleeding*. The patient's blood cells are tagged with ^{51}Cr and the agent under test is administered. If GI bleeding occurs, there is an increase in ^{51}Cr content of fecal samples as a result of blood loss into the lumen of the GI tract.

Measurement of the absorption of radioactive iron (^{59}Fe), its tissue distribution (liver, spleen, precordium, sacral bone marrow), plasma elimination and urinary excretion establish various *ferrokinetic parameters*. Iron is absorbed to the greatest extent as the ferrous salt in the upper small intestine. Absorption is decreased in iron overload, erythropoiesis and various malignant, inflammatory or infectious diseases. Iron is transported in plasma bound to transferrin, a specific iron-binding protein. Alterations in plasma iron and iron-binding capacity are seen in pregnancy, thalassemia major and iron deficiency (hypochromic) anemia. Iron is stored in the liver, bone marrow, skeletal muscle and spleen as ferritin and hemosiderin. The daily turnover of iron is about 35 mg, primarily from an "erythropoietic labile pool" in bone marrow.

Hemosiderosis is simply an increase in iron storage, whereas *hemochromatosis* denotes increased iron storage with associated tissue damage. Both of these states can result from oral or parenteral medicinal/transfusion iron overload. Iron excretion is limited and occurs by desquamation of iron-containing cells from the bowel, skin and urinary tract.

Iron-deficiency anemia is a symptom and not a disease. Treatment is based on evaluation of ferrokinetic parameters, correction of hemoglobin and tissue-iron deficiency and recognition of the underlying cause (eg, chronic blood loss).

Blood Coagulation—*Hemostasis*, the arrest of blood flow from a vessel, is regulated by extravascular (muscle, skin and subcutaneous tissue), vascular (blood vessels) and intravascular (platelet-adhesion, clot-retraction and blood-coagulation) mechanisms. The following discussion will be limited to those processes related to the blood-coagulation mechanism. When blood is allowed to clot, the free-flowing liquid is converted into a firm cell clot surrounded by serum. If an anticoagulant is added to blood, coagulation does not occur and the blood cells are suspended in a liquid phase —plasma. The clotting mechanism involves three stages: the formation of plasma *thromboplastin*, the conversion of *prothrombin* to *thrombin* and the conversion of *fibrinogen* to *fibrin*.

The International Committee on Nomenclature of Blood Clotting Factors has numerically designated the blood-coagulation factors (Table II). Fibrinogen and Factors V and VIII are absent in normal blood serum as a result of the clotting process. The absorption characteristics of certain blood-coagulation factors on calcium phosphate or barium sulfate are used in the differential analysis of specific factors. The interaction of coagulation factors may be initiated through either the intrinsic or extrinsic pathways. In the intrinsic system all the factors are present in the blood, while the extrinsic system is activated by the release of tissue thromboplastin. Figure 28-3 shows the activities of both pathways to form a stabilized fibrin clot.

In Stage 1 of the coagulation process, the contact of injured tissue with blood results in the activation of Factor XII, which reacts with calcium, PTA, PTC, AHG and Factors III, V and X to yield intrinsic or blood thromboplastin. This stage normally is completed in 3 to 5 min. Extrinsic or tissue thromboplastin is formed rapidly (<12 sec) in various tissues in the body such as lung and brain in the presence of calcium and Factors V, VII and X.

Table II—Blood-Coagulation Factors

Factor	Synonym
I	Fibrinogen
II	Prothrombin
III	Thromboplastin (tissue)
IV	Calcium
V	Labile factor, proaccelerin, Ac globulin
VI	Accelerin
VII	Stable factor, proconvertin, serum prothrombin conversion accelerator (SPCA)
VIII	Antihemophilic globulin (AHG)
IX	Christmas factor, plasma thromboplastin component (PTC)
X	Stuart-Prower factor
XI	Plasma thromboplastin antecedent (PTA)
XII	Hageman factor
XIII	Fibrin-stabilizing factor (FSF)

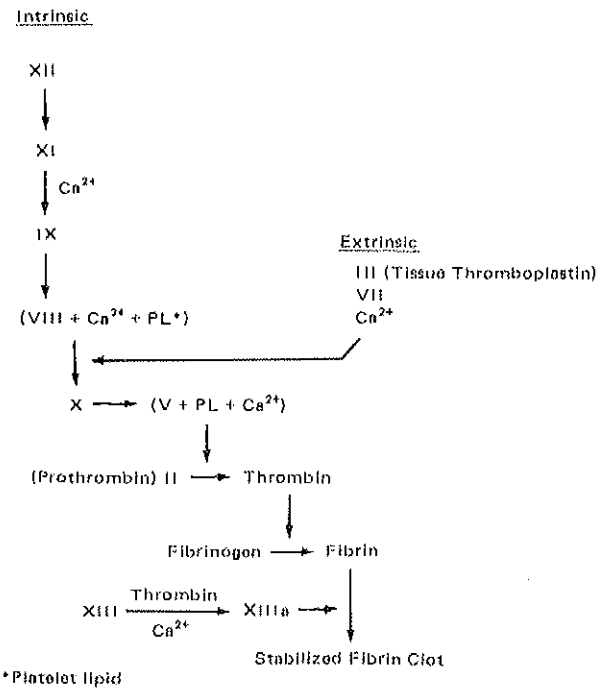


Fig 28-3. Blood Coagulation Process.

In Stage 2, thromboplastin catalyzes the conversion of prothrombin to thrombin (8 to 15 sec) in the presence of Factors V, VII, X and calcium.

In Stage 3, the thrombin rapidly converts fibrinogen into fibrin, which then forms a network of fibers that traps red cells and thus forms the blood clot.

Although the exact nature of the enzymatic sequences in the coagulation process is not clear, it is definitely a biological amplification process starting from the small reaction of tissue contact to rapid conversion of fibrinogen to fibrin.

Blood contains natural inhibitors of coagulation such as antithrombin, heparin and antithromboplastin, which can prevent a particular reaction in the coagulation sequence. The dissolution of blood clots occurs by the action of the blood proteolytic enzyme—plasmin or fibrinolysin. Plasmin is formed from its precursor, plasminogen, after activation by tissue and body fluids or substances of bacterial origin (streptokinase).

The routine tests performed in the coagulation laboratory are indices of vascular function (vascular phase and platelet adhesion) or intrinsic clotting mechanisms. Determinations of *bleeding time* and *capillary fragility* provide esti-

mates of blood coagulation in the presence of platelets and tissue or vascular factors. In the Ivy method for determination of *capillary bleeding time*, a blood pressure cuff is placed on the forearm and inflated to 40 torr; a puncture wound is made and the time required for bleeding to stop is noted. *Bleeding time* is a screening test for disorders of platelet function or vascular defects but is usually normal in coagulation disorders. The test is useful in the differential diagnosis of Von Willebrand's (reduced factor VIII, with a normal bleeding time) disease from mild hemophilia. The normal bleeding time, as determined by this method is 1 to 9 min. Dextran, pantothenyl alcohol, and derivatives, penicillin G, nonsteroidal anti-inflammatory drugs and streptokinase-streptodornase may cause a prolonged bleeding time. The *Simplate II* (General Diagnostics Div, *Warner Lambert*) is a standardized, disposable, springloaded bleeding-time device for platelet function testing. It uses two blades that are released automatically to produce two uniform incisions 6 mm long × 1 mm deep, making the procedure reliable and reproducible.

The *capillary fragility* or *tourniquet test* is based on the incidence of petechiae (small red marks) formation produced by an inflated blood pressure cuff over a 5-min period. Normally, a few tiny petechiae may appear. The most common cause of abnormalities in vascular-function and platelet-adhesion tests is thrombocytopenia.

An analysis of the *intrinsic coagulation mechanism* is concerned with the determination of the levels of the specific clotting factors in whole blood. In preliminary studies of a suspected hemorrhagic disorder, determinations of *coagulation time, clot retraction, platelet count, bleeding time* and *capillary fragility* usually are performed.

In the Lee-White procedure, the coagulation time of whole blood is determined in regular or siliconed tubes. Normal values are 8.5 to 15 min in glass and 19 to 60 min in siliconed tubes. Anticoagulants and tetracyclines may cause increased times while corticosteroids and epinephrine cause decreased values. The siliconization of glassware prevents platelet aggregation and thus, delays coagulation. The samples used in the analysis of coagulation time are then inspected at 0.5, 1, 2, 4 and 24 hr after clotting to determine the time required for the various phases of clot retraction. The tubes also are observed for evidence of clot lysis or dissolution. The clot normally will start to retract in 30 min, completely retract within 24 hr and show no evidence of lysis over a 72-hr period. Prolonged coagulation times are associated with hemophilia, hypofibrinogenemia and Factor IX deficiency. Abnormalities in any of these tests indicate the requirements for further coagulation studies.

The *prothrombin time test* is a measure of the levels of all coagulation factors, except III, IV and VII, and is an index of the capacity of plasma to form thrombin. In the "One Stage" test, the plasma sample is mixed with calcium chloride and tissue thromboplastin, and the time required for fibrin-clot formation is determined. Results are compared with a normal plasma control, and the prothrombin time is reported either in seconds or as the percent of prothrombin calculated from a standard activity curve. Correction studies using normal serum, adsorbed normal plasma or whole normal plasma added to test serum indicate deficiencies of Factors VII and X, Factor V and Factor II, respectively. If none of these additives shorten the prothrombin time, a circulating anticoagulant problem can be suspected.

A modification of this technique (the *prothrombin-proconvertin procedure*) using a 1:10 dilution of both patient and control plasma in the presence of prothrombin-free plasma as a source of Factors I and V, is a more sensitive index of specific deficiencies in prothrombin, Factor VII, IX and X.

Owren's *thrombotest*, as performed on whole blood, is sensitive to changes in both extravascular and intravascular clotting mechanisms, including Factor IX. The dosage of anticoagulant drugs, such as dicumarol, is adjusted in accordance with prothrombin-time determinations; patients are maintained usually within a therapeutic range of 20 to 40% of prothrombin activity (normal range, 80 to 130%). Reduced prothrombin levels, with prolonged prothrombin times, are observed in vitamin K deficiency, hemorrhagic disease of the newborn, excessive anticoagulant therapy, liver and biliary disease. The interaction of other drugs with anticoagulants may cause increased prothrombin times. Drugs such as salicylates, phenylbutazone, oxyphenbutazone, indomethacin and some sulfonamides increase the amount of active anticoagulant activity. Other drugs decrease the amount of vitamin K produced by gut bacteria which include chloramphenicol, kanamycin, neomycin, streptomycin and the sulfonamides.

The *prothrombin consumption test* is an index of the efficiency of conversion of prothrombin to thrombin in the coagulation process. The blood sample is allowed to clot under standardized conditions and then the quantity of prothrombin complex removed in the serum is determined in the presence of extrinsic fibrinogen. At least 80% of the prothrombin is consumed normally. Reduced consumption of prothrombin (<80%) is observed in coagulation deficiencies (hemophilia) related to thromboplastin generation.

Other types of coagulation tests detect deficiencies in *thromboplastin generation mechanism*. The *thromboplastin generation time test* (TGT) provides a means of detecting specific deficiencies of Factors V, VIII, IX, X, XI or XII. In the initial phase of this procedure the clotting time of the patient's adsorbed plasma is determined in the presence of a standardized platelet factor reagent, calcium chloride, plasma substrate reagent (Factors I, II and V) and the patient's serum. If the clotting time is abnormal (>16 sec), further tests are performed with the patient's plasma or serum. The adsorption of the plasma sample on barium sulfate removes Factors II, VII, IX and X and facilitates differentiation of a Factor IX to X from V to VIII deficiency in the thromboplastin-generation mechanism. Thromboplastin generation is reduced in hemophilia and thrombocytopenia.

The *activated partial thromboplastin time test* (PTT) is based on the observation that hemophilic plasma has a normal clotting time in the presence of a complete thromboplastin (extrinsic-saline extract of brain tissue), as used in prothrombin determinations, but will give a markedly prolonged clotting time with an incomplete thromboplastin (cephalin). Cephalin is a thromboplastic, ether-soluble phospholipid factor with platelet-like activity. In this test the clotting time of the patient's plasma is determined in the presence of calcium chloride and activated cephalin. This test is used primarily to detect deficiencies in Stage 1 of the coagulation mechanism and is rather sensitive to changes in Factors VIII and IX, as seen in classical hemophilia and Factor IX deficiency (Hemophilia B or Christmas disease).

In Stage 3 of the coagulation process, the presence of adequate levels of fibrinogen and thrombin is critical. *Fibrinogen levels* are analyzed semiquantitatively by determining the clotting time of a diluted plasma sample in the presence of extrinsic thromboplastin. This test is basically independent of prothrombin levels. Fibrinogen concentrations of 125 mg% or greater are adequate; deficiencies (hypofibrinogenemia) have been observed in liver disease, carcinomatosis and in certain complications of pregnancy.

Increased levels of *fibrinogen degradation products* (FDP) have been demonstrated in serum due to primary activation of the fibrinolytic system (pathological fibrinolysis) or by secondary activation following increased blood clotting (disseminated intravascular coagulation). Fibrinogen (mol wt 3.4×10^6) is degraded sequentially to fragments

X, Y, D and E with mol wts of 2.7, 1.65, 0.85 and 0.55 × 10^5, respectively. Fragments X and Y are more potent anticoagulants than fragments D and E and are responsible for hemorrhagic states in defibrination. Complexes between fibrin monomer, fragment X and other FDP interfere with thromboplastin generation and platelet formation. FDP can be measured by immunological techniques involving latex agglutination of particles sensitized with specific antibodies to FDP or by a hemaglutination-inhibition test. The normal level of serum FDP is 4.9 ± 2.8 μg/mL. Increased levels are seen in acute myocardial infarction, menstruation, complications of pregnancy, hypoxic newborns, malignancy and renal disease.

Deficiencies in the clotting mechanisms usually can be corrected partially and temporarily by transfusion of normal blood or plasma. When this fails, the presence of *circulating anticoagulants* (antithrombin, antithromboplastins, heparin) must be considered. Heparin acts indirectly by means of antithrombin III, which neutralizes several activated clotting factors (XIIa, activated Fletcher factor, XIa, IXa, Xa, IIa and XIIIa). The pharmacological effect of an oral anticoagulant is the inhibition of blood clotting by interfering with vitamin K-dependent clotting factors II, VII, IX and X. Circulating anticoagulants are detected by determining the effect of normal plasma on the clotting time (*recalcification time*) of the patient's oxalated plasma in the presence of calcium chloride. If the addition of the normal plasma does not shorten the prolonged recalcification time, a circulating anticoagulant state can be reported.

Since the end-point of all coagulation tests is the conversion of fibrinogen to fibrin, it is vital that the analyst rigidly standardize his concepts of fibrin formation in visual recording procedures. The use of mechanical instrumentation in the detection of clot formation significantly has increased the standardization, accuracy and reproducibility of coagulation procedures. These instruments measure and record the process of fibrin formation via increased turbidity (coagulogram or photometric clot detection) or changes in electrical conductance in the reaction mixtures. As well as performing routine laboratory tests simultaneously or sequentially, updated systems can run Fibrinogen and Factor assays achieving rapid throughput and accuracy. New performance features are available with many of the automated coagulation instruments. These include monitoring temperature zones, digital displays of the individual clotting times, automatic dilutions or patients samples and programmable parameters for testing flexibility.

Hemophilia is a classic deficiency of antihemophilic globulin (AHG), Christmas disease of PTC and Hageman trait of Factor XII. Hereditary or acquired deficiencies of Factors II, V, VII, X and XI also are associated with disease states. The process of blood coagulation, analysis of coagulation factors and interpretation of results comprise a highly complex system. The coagulation laboratory and the physician function together in the diagnosis and treatment of coagulation-deficiency diseases.

Blood-Bank Technology

Blood-bank technology in the modern laboratory is part of the blood-transfusion service. As whole blood for transfusion and its components are biologically active therapeutic substances, a complete analysis of their chemical and biological characteristics is vital to the assurance of successful therapeutic effects. The transfusion service is responsible for:

1. Receiving and examining of the donor.
2. Collecting, processing and storing the blood.
3. Typing of recipient and donor for ABO and Rh blood-group factors.
4. Compatibility (cross-matching) testing before transfusion.
5. Issuing of blood for transfusion and extracorporeal circulation.
6. Evaluating transfusion complications.
7. Performing of special serological tests pertinent to blood groups and other factors.

In this section a discussion of pertinent factors related to the various phases of the transfusion service will be presented.

Receiving and Examining of the Donor—A complete registry[4] of prospective donors should be maintained, with specific reference to age, sex, weight, address, occupation and telephone number. Computerized blood banking has increased the efficiency of this service. Donors should preferably be between the ages of 21 and 60 and should weigh no less than 110 lb. The donor may be rejected on the basis of previous or active incidence of certain microbial diseases (recurrent malaria, syphilis, infectious or homologous serum hepatitis, tuberculosis), bleeding abnormalities, convulsions, allergic syndromes, skin or heart diseases, diabetes, alcohol or drug addiction, pregnancy, cancer, recent immunization with live vaccine product, acquired immune deficiency syndrome (AIDS) or blood-pressure abnormalities (acceptable blood pressure: between 100/50 and 200/100; pulse rate: 60 to 12/min). The screening of blood for exposure to human immunodeficiency virus (HIV) is crucial to reducing the risk of infection from transfusion. ELISA (enzyme-linked immunosorbent assay) screening tests for the detection of antibodies against HIV are available from manufacturers. More sensitive tests are being developed to detect viral DNA in body fluids.

A period of at least 8 weeks should have elapsed since blood was withdrawn and the blood hemoglobin level should be 12.5 to 13.5 g% or greater. Serum bilirubin and transaminase levels also should be evaluated in donors with previous incidence of jaundice.

Collecting, Processing and Storing the Blood—A tourniquet is applied to the arm of the donor to occlude the venous return, the skin area is sterilized and the blood is collected by venipuncture (phlebotomy). NIH Formula A or B [ACD(Acid–Citrate-Dextrose) or ACD-phosphate] solutions are used as anticoagulants in the sterile blood-collecting containers. Evacuated containers may be of regular or siliconed glass; collapsible plastic containers offer many advantages in donation, blood-banking and transfusion procedures.

The preservation of the red cells in blood is improved by the complete removal of trapped air in the blood-collection apparatus, rapid cooling after collection and storage at 4°. Properly collected whole blood is usually stable for 21 days at 1 to 6°. The deterioration of whole blood is related to increased cellular fragility (increased plasma K^+) and decreased glucose utilization. Blood which is used for correction of any bleeding tendency or clotting defect should be as fresh as possible. Leukocytes, platelets and Factors V and VIII deteriorate in stored plasma or whole blood.

ABO Blood-Group Classification[5]—Human red cells can be classified into various groups or types on the basis of reactivity of certain blood factors (*agglutinogens*) located on the erythrocyte membrane. The Landsteiner system (Table III) for the four blood groups is based on the presence or absence of either A or B agglutinogen on the cell surface (Group A, B, AB or O, respectively).

Serum does not contain the antibody (*agglutinin*-IgM type) for the antigen present in an individual's own red cells, but does contain the isoagglutinin (eg, anti-B in blood group A) due to exposure, early in life, to bacterial and plant antigens similar in structure to the A-B antigens. The clumping or agglutination of the red cells by reaction of agglutinogen with agglutinin is used in blood-grouping techniques. In certain instances hemolysin antibodies, present

Table III—Blood-Group Systems

Blood Group	Agglutinogen in Cell	Agglutinin in Serum	Reaction[a] with Anti-A Serum	Reaction[a] with Anti-B Serum	Frequency (%) in Caucasians
A	A	Anti-B	+	−	41
B	B	Anti-A-A_1	−	+	10
AB	AB	None	+	+	4
O	None	Anti-A and B	−	−	45

[a] Agglutination.

in serum containing anti-A or anti-B agglutinins, cause the disruption of cells and release of hemoglobin (hemolysis).

Human blood cells are grouped by two separate reactions: cellular or "front" grouping and serum or "reverse" grouping. The blood group ordinarily is determined by testing an individual's red cells with standardized anti-A or anti-B serum (certified by the Div of Biological Standards, NIH). Confirmation of the blood group (reverse typing) is accomplished by an analysis of an individual's agglutinin titer. In this procedure the individual's serum is heated at 56° for 10 min to destroy hemolysins, and then mixed with known Subgroup A_1 or B_1 human red (Rh-negative) cells in the agglutination test. These two tests should be in agreement prior to the release of blood for transfusion.

Although human blood cells of Group B react uniformly with Anti-B serum, Group A and AB cells show a wide range of reactivity with Anti-A or Anti-A_1B serum. Blood-group A may be further categorized into Subgroups A_1, $A_{int.}$, A_2, A_3, A_0 and A_x on the basis of the reaction with absorbed Anti-A, Anti-A_1-lectin, Anti-H-lectin, Anti-$A_{1,2}$ and Anti-AB serum and the presence of Anti-A_1 in the serum. Certain Group O individuals possess anti-H in their serum and are further subcategorized into the Bombay or O_h phenotype. Tests for A, B and H in saliva can establish the genotype of an individual, ie, A and H in saliva of blood-group A, B and H in B, H and O and A, B, H in AB. This is helpful in cases of poorly developed red-cell antigens or in the loss of cellular antigen in some patients with leukemia.

As the human blood cell contains many antigens with rather complex biochemical and immunochemical properties, the blood factors have been classified further into various subsystems. The Kell (K), Lutheran (Lu), Lewis (Le), Duffy (Fy), Kidd (Jk), MNS, Sutter (Js), Diego (Di) and P blood-factor systems are based on the detection of a specific antigen on or within the red cell by means of antibody (*isohemagglutinin*) reactions with specific antisera or panels of reagent red cells. Some of these factors (eg, Kidd, Kell and Lewis) have been involved in transfusion reactions.

The Rh-Hr System and Antihuman Globulin Test—The presence or absence of *Rh_0 antigen* in human blood is of prime importance in transfusion reactions, paternity disputes and isosensitization phenomena. There are eight blood Rh phenotypes which are determined by their reaction with three specific serum agglutinins (Anti-Rh_0, Anti-rh′ and Anti-rh″): rh, rh′, rh″, rh′rh″, Rh_0, Rh_0′, Rh_0″ and Rh_0′Rh_0″. The rh groups do not contain the Rh_0 factor on the cell surface and are designated "Rh-negative." The terminology of the Wiener system (Rh, rh) is comparable to the Fisher-Race (CDE) as follows: rh′(C), Rh_0(D), rh″(E). The Rosenfeld system uses a numerical classification: RH1 = Rh_0.

The absence of the Rh antigen in about 15% of the population does not preclude the presence of other factors; the use of specific antisera (Anti-hr′ and Anti-hr″) has demonstrated the existence of the Hr factors (Hr_0, hr′, hr″). For example, the Rh-negative cell (rh″) possesses rh″hr′Hr_0 antigens. The antigen Rh_0(D) is the most potent immunogen of all the Rh antigens.

The Rh antibodies are either *saline agglutinins* (complete) or "*blocking*" *antibodies* (incomplete). The latter are of the IgG type. They are used in Rh testing procedures and are produced more commonly, and in higher titer, in the human isosensitization or autoantibody reactions. They will not agglutinate saline suspensions of normal Rh-positive red cells except in the presence of a high concentration of albumin, serum or conglutinin (AB serum with albumin) at a temperature of 35 to 37°.

In routine Rh testing procedures, a sample of blood (oxalated or heparinized) or a suspension of cells in serum or albumin is mixed with Anti-Rh_0 serum on a slide or in a tube at 37 to 47°. The presence of clumping indicates that the blood possesses Rh_0 antigen. Confirmation of an Rh-negative test may be performed by retesting with Anti-rh′/Rh_0rh″ serum.

In Rh testing procedures, red cells from patients with acquired hemolytic anemia are partially coated with human autoantibody, and cells from erythroblastic infants are coated with maternal antibody globulins and may be clumped falsely by Rh typing serum containing a high protein concentration, or may appear to be Rh-positive in the saline-cell suspension test. Demonstration of anti-Rh_0(D) in an eluate from these antibody-coated cells can help to establish true Rh type.

Anti-Rh antibodies are not normally present in human serum; they may be acquired via isosensitization. The transfusion of Rh-positive blood to an Rh-negative recipient, or transfer of cells of Rh-positive fetus through the placental barrier to the Rh-negative mother, will result in formation of antibodies to Rh agglutinogens not present in the cells of the recipient or mother, respectively.

Hemolytic blood-transfusion reactions and hemolytic disease of the newborn (erythroblastosis fetalis) involve *isosensitization phenomena* usually related to the Rh_0 antigen. Hr and ABO antigens also can be responsible for hemolytic disease of the newborn. If an expectant mother is Rh-negative and the father is Rh-positive, the Rh genotype of the father should be determined. If the father is homozygous, the erythrocytes will contain a pair of Rh_0 factors and the offspring will inherit the Rh_0 factor; if he is heterozygous, one Rh_0 and one Hr_0 factor will be present and his offspring may or may not inherit the factor.

If the fetus is Rh-positive, the mother may be sensitized to the Rh antigen and in subsequent pregnancies the development of high titers of Anti-Rh_0 antibodies will result in hemolytic disease of the fetus. These antibodies enter the fetal circulation via the placental barrier, coat the red cells of the fetus and cause excessive erythrocyte destruction, hyperbilirubinemia and associated potential for brain damage, hydrops fetalis (edema) and congenital anemia of the newborn. This Rh disease can be avoided now by proper therapeutic use of Rh_0(D) Human Immune Globulin (RhoGAM, *Ortho*) to prevent the postpartum formation of active antibodies in the Rh_0(D)-negative, D^u-negative mother who has delivered an Rh_0(D)-positive or D^u-positive infant.

The *Coombs' antiglobulin test* is a method of detecting the blocking-type antibodies, globulins and complement which are attached to red-cell antigens in isosensitization phenomena.

In the "direct" test procedure, a saline suspension of washed red cells is mixed with anti-human gamma globulin antiserum and agglutination is indicative of the combination of human antibody with antigen on the red cell, eg, maternal incomplete isoantibody on infant's red cells in hemolytic disease of the newborn, autoimmune, drug-induced, alloantibody-induced hemolytic anemia and after transfusion of incompatible red cells.

An "indirect" procedure is used to demonstrate the presence of blocking antibody in the serum of pregnant Rh-

negative women and in transfusion reactions. In this procedure the patient's serum is incubated with a suspension of Group O Rh-positive red cells; the cells are washed and then antihuman globulin antiserum is added to detect the coating of the red cells with antibody globulin from the patient's serum by agglutination phenomena. If agglutination occurs in the first part of the procedure, a saline agglutinin is also present. Anticomplement sera (anti-nongammaglobulin antiserum) are used to detect reactions involving anti-JK.

The Du allele is a clinically important variant of the Rh_0 factor and usually associated with rh′(C) and rh″(E). Individuals with this factor are considered Rh-positive, and the red cells fail to react with anti-Rh_0 in the saline-tube method but reacts with incomplete anti-Rh_0(D) by other slide or tube techniques. Rh-negative donors should be tested for Du factor. If positive, their blood must only be given to Rh-positive recipients.

Drug-Related Problems—Hematological abnormalities may be caused by the administration of drugs which can cause a positive direct antiglobulin test and immune hemolytic anemia, eg, cephaloridine, cephalothin (*Keflin*), methyldopa (*Aldomet*), penicillin, L-dopa, quinidine, phenacetin and insulin.

Compatibility Testing—Cross-matching procedures are designed to detect incompatibilities in the blood of donors and recipient. The test is designed to prevent transfusion reaction and assure maximum benefit to the patient. Although erroneous ABO grouping usually will result in an incompatible cross match, no such protection exists in the Rh system. An incorrectly typed Rh-positive donor blood can result in primary immunization to Rh_0(D) antigen if transfused to an Rh-negative recipient. For each transfusion, a *major* and *minor cross match* should be performed.

In the *major cross match* (1) a saline suspension of the donor's cells is mixed with the recipient's serum and (2) the donor's cells are suspended in recipient's serum or in serum with added albumin. The saline cross match is an additional check on the ABO typing and may detect incompatibilities caused by antibodies to M, N, S, P and Lu subgroups. The high-protein or albumin cross match can demonstrate antibodies in the Rh system. The presence of agglutination or hemolysis indicates incompatibility.

The *minor* cross match includes the donor's serum and the recipient's cells, and is useful as a check of the ABO typing and an indication of the possibility of transfusion reactions caused by a rare antigen on the recipient's cells or uncommon antibodies directed against an antigen in the serum of the donor. The minor cross match has been replaced in many instances with screening of the donor's serum against a panel or pool of red cells of known antigenicity.

The *indirect antihuman globulin* procedure also must be performed with the recipient's serum and donor's cells with and without albumin (major side) and may be tested with the donor's serum and recipient's cells (minor side). The use of proteolytic enzymes (bromelain) enhances the agglutination of red cells by low-titer or weakly reacting Rh-Hr antibodies, probably by removing sialic acid residues on the RBC surface. The red cells used in the indirect Coombs test are treated with the enzyme prior to absorption of antibodies and addition of antiglobulin reagent.

The usual cross-matching techniques involve (1) a room-temperature or 30° procedure, preferably with the addition of albumin, (2) a high-protein procedure and (3) an antiglobulin procedure.

The presence of nonspecific *autoantibodies*, *cold agglutinins* and *bacteriogenic agglutination* sometimes complicates the cross-matching procedure. If the recipient's serum reacts more strongly with his own cells than with the donor's, autoantibodies should be suspected. Cold agglutinins usually will agglutinate all blood, regardless of type, at low temperatures, but will not react at 37°. Agglutination as a result of bacterial contamination of blood is called panagglutination.

Hepatitis Testing—Posttransfusion hepatitis is associated with the transmission of virus-like particles referred to as *Australia or serum hepatitis antigen or the hepatitis associated antigen (HAA)*. All donor blood must be tested for the presence of HAA. Agar gel diffusion (AGD), counterelectrophoresis (CEP), complement fixation (CF) and rheophoresis procedures can be used.[6] The rheophoresis procedure uses a modified gel-diffusion technique for the detection of HAA by precipitin-type reaction with HAA antibody. It offers the sensitivity of CEP and CF procedures with the simplicity of the AGD procedure. Other tests for HAA are based on radioimmunoassay (RIA) technique for detection of antigen by hemagglutination (HA) or HA-inhibition for the presence of HAA antibody. In the RIA technique, the donor's serum is added to a test tube coated with HAA antibody (solid RIA). If the serum contains HAA, it will bind to the antibody. ^{125}I-HAA is then added to the tube. If the antibody binding site is occupied previously with HAA from the donor's serum, ^{125}I-HAA will not bind and the determination of ^{125}I bound versus free is an index of HAA content of the donor's serum.

Issuing of Blood and Evaluating Transfusion Reactions—Whole-blood, red-cell or leukocyte suspensions, plasma, platelet-rich plasma, platelet concentrates, leukocyte-poor blood, AHF, factor IX complex, plasma protein fractions and RhoGAM are products of the transfusion service.[7] Transfusion reactions are related to antibody phenomena or disease transmission. The hemolytic reaction resulting from the transfusion of incompatible cells is the most serious problem. The transfusion of microbially contaminated blood can result in a pyrogenic reaction or transmission of infectious diseases, such as malaria, syphilis or hepatitis. Allergic reactions (urticaria, asthmatic seizures), circulatory overload, embolic complications (blood clot, air emboli) also may be encountered. Leukocyte and platelet antibodies develop in repeat transfusions and in transplantation patients. The transfusion service is an integral unit in evaluating such complications.

Techniques of Analysis

This section will describe the principles of the procedures used in the analyses of various substances in blood, plasma or urine. Examples of the significance of such tests in clinical diagnosis will be presented. For a complete description of the physiological and pharmacological aspects of these blood constituents, see the *Bibliography*.

Instrumentation—The development of instrumentation has accelerated progress in clinical chemistry. An excellent review of the principles and applications in clinical chemistry of automation, atomic-absorption spectroscopy, ultraviolet and visible spectrophotometry, fluorimetry, phosphorimetry, infrared and Raman spectroscopy, microwave and radiowave spectroscopy and nucleonics was prepared by Broughton and Dawson.[8] Quality-control techniques are a vital part of any clinical laboratory. Standard reference materials,[9,10] standardization of quantities and units[11] and continual evaluation of precision and accuracy of various determinations[12] are incorporated into procedures of all reliable clinical laboratories. The manufacture of certified standards and reagents and the certification of clinical chemists and clinical laboratories are under the supervision of either the FDA, NIH, Pharmaceutical Manufacturers Association (PMA), American Association of Clinical Chemists, the College of American Pathologists and the National Committee for Clinical Laboratory Standards (NCCLS).

Interaction of Drugs with Clinical Laboratory Tests—Drugs may interfere with the interpretation of laboratory tests by three classes of mechanisms:

I. *Chemical or biochemical* interference due to reaction of a drug or its metabolite in biological fluids with test reagents in analytical procedures.
II. *Pharmacological* interference due to normal drug-induced alterations in various physiological parameters.
III. *Toxicological* interference as a consequence of the toxicity of a drug.

Examples of Class I interference include false-positive urine glucose results due to the reducing properties of drugs or metabolites such as ascorbic acid, *p*-aminosalicylic acid, tetracycline, cephaloridine and levodopa, which are excreted in urine. Spironolactone will result in an elevation of certain urinary ketosteroids through cross-reaction of the drug in the analytical procedure.

Examples of Class II interference include the decrease in serum-potassium levels in patients receiving thiazide diuretics, the alteration in serum uric acid with probenecid and the elevation in various plasma proteins and thyroid function tests with estrogen-progesterone combinations. Drug-drug interaction also can result in changes in these parameters. Guanethidine enhances the effect of the coumarin anticoagulants. Barbiturates induce hepatic microsomal enzyme synthesis and subsequently increase the metabolism and decrease the therapeutic effect of drugs, such as warfarin, even after these drugs are terminated.

Examples of Class III interference include changes in liver- and kidney-function tests and hematological parameters (anemia, agranulocytosis, leukopenia) due to drug-induced toxicity and positive LE and ANA tests due to a "lupus-like" syndrome induced by hydralazine.

It is beyond the scope of this chapter to include a complete listing of drug interactions in laboratory tests. The reader is referred to an annual, readily available, computerized review of the effect of normal therapeutic drug doses, as well as overdoses, on clinical laboratory tests[13] and to other review articles.[14]

Blood

Collection and Preparation for Chemical Analysis—Using aseptic technique, a blood sample is obtained by venipuncture and usually placed in evacuated glass tubes. The choice of anticoagulant, type of specimen, stability of test component and use of preservatives depends on the type of analysis requested and the specific analytical procedure involved. If serum is desired, the blood sample is allowed to clot and the serum is separated by centrifugation. When whole blood or plasma is to be used in the analysis, an anticoagulant is added to the collecting tube.

The following concentrations of specific anticoagulants are used routinely per 10 mL blood; lithium, potassium or sodium oxalate (15 to 25 mg), sodium citrate (40 to 60 mg), heparin sodium (2 mg), disodium or tripotassium ethylenediaminetetraacetate (EDTA-Na_2, 10 to 30 mg) or ACD-Formula B solution (1.0 mL).

Heparin prevents blood coagulation by inhibiting the thrombin-catalyzed conversion of fibrinogen to fibrin. The other anticoagulants either precipitate blood calcium or convert ionized calcium into a nonionized (chelated) form which cannot function in the coagulation reaction. Heparin and EDTA do not alter the cellular elements of blood significantly. Sodium fluoride and thymol are used as preservatives or enzyme inhibitors to prevent the deterioration of various substances in the blood sample, eg, glucose → lactic acid. Preservatives and anticoagulants can interfere with some enzyme tests. Serum usually is used for these procedures.

The separation of plasma or serum, and chemical analysis, usually are performed as soon as possible after the collection of the sample. The addition of polystyrene granules to the blood sample prior to centrifugation facilitates the isolation of serum or plasma. Hemolysis interferes with analytical procedures for bilirubin, albumin, nonprotein nitrogens, pH, phosphorus, potassium and various enzymes. The serum also should be observed for presence of lipemia. Changes in the ratio of CO_2, chloride and electrolytes in cells and plasma, glycolytic conversion of glucose to lactic acid, hydrolysis of ester phosphate to free inorganic phosphate, bacterial conversion of urea to ammonia and conversion of pyruvate to lactate are examples of changes that can occur in contaminated, improperly preserved or unrefrigerated blood specimens.

The first stage in many of the chemical determinations is the removal of blood protein and preparation of *protein-free blood filtrate*. The protein is precipitated with tungstic acid, trichloroacetic acid, zinc hydroxide or organic solvents, such as alcohol and acetone, and then filtered or centrifuged to remove the protein coagulum. Tungstic acid precipitation is performed by mixing 1 volume of blood or 2 volumes of plasma with 9 volumes of stabilized tungstic acid reagent. The filtrate obtained in this procedure should be in the pH range of 3.0 to 5.1 to assure the adequate removal of proteins (<2 mg% in filtrate).

The Somogyi filtrate is prepared by mixing 1 volume of blood with 5 volumes of water, 2 volumes of 5% zinc sulfate and 2 volumes of 0.3 *N* barium hydroxide. The barium sulfate is precipitated and the zinc hydroxide formed in the reaction precipitates the blood proteins. Trichloroacetic acid (10%), in a ratio of 9:1 with blood, yields greater volumes of filtrate due to a more complete formation of protein agglomerates.

Blood Glucose—Methods for determining blood glucose are based on the use of glucose as a reducing agent or on the enzymatic oxidation of glucose to gluconic acid. In the Folin-Wu technique, glucose is determined in a protein-free blood filtrate by reduction of alkaline cupric sulfate and subsequent reaction with phosphomolybdic or arsenomolybdic acid reagent to form a blue complex which can be estimated colorimetrically. The Nelson-Somogyi method uses a protein-free blood filtrate prepared with zinc hydroxide to remove most of the interfering reducing substances.

The presence of a terminal aldehyde in the glucose molecule is the basis of a colorimetric determination with phenolic hydroxyl reagents (phenol in aqueous methyl salicylate or phosphorylated 1,3-dihydroxybenzene) in the presence of strong sulfuric acid and heat.

The *o*-toluidine procedure is a color reaction specific for hexoses—glucose, mannose and galactose. Since aldohexoses other than glucose are normally present in very small concentrations, results obtained by this method approach the true value of glucose. *o*-Toluidine is condensed with glucose in glacial acetic acid to yield a green chromogen by forming an equilibrium mixture of a glycosylamine and Schiff base.

In the preceding techniques, interfering substances such as lactose, galactose and glutathione are measured and the value is reported in the nonspecific term "sugar." Enzymatic determination with glucose oxidase is the only test specific for blood glucose. Blood glucose is converted to gluconic acid and hydrogen peroxide by glucose oxidase; the peroxide is then estimated by iodimetric procedures or by oxidation of a chromogen (*o*-dianisidine or 2,2′-azino[diethylbenzothiazolinesulfonic acid]) in the presence of a peroxidase to form a colored product. Drugs which cause a slight increase in glucose values include ACTH, corticosteroids, D-thyroxine, diazoxide, epinephrine, estrogens, indomethacin, oral contraceptives, lithium carbonate, phenothiazones, phenytoin, thiabendazole and diuretics. Drug interferences with *o*-toluidine methods, which cause a slight increase, include

ascorbic acid, dextran, fructose, galactose, mannose, ribose, xylose and bilirubin.

Another enzymatic procedure uses the hexokinase-catalyzed conversion of glucose to glucose 6-phosphate (G6P), and then to 6-phosphogluconate and NADPH in the presence of NADP and G6P dehydrogenase. The NADPH thus formed is equivalent to the amount of glucose present and is estimated spectrometrically at 340 or 366 nm.

Normal fasting blood-sugar values for adults are 80 to 120 mg/100 mL; true glucose is 65 to 100 mg/100 mL. When the blood-sugar values exceeds 120 (hyperglycemia), diabetes mellitus should be suspected and can be confirmed by evidence of diminished carbohydrate tolerance. The effect of ingested carbohydrate on blood sugar can be determined by the *glucose tolerance test;* 100 g of glucose (1.75 g/kg) in water or a flavored beverage, is administered orally and glucose determinations are performed on blood and urine samples at hourly intervals for 3 hours. Values above 160 at 1 hr and 110 at 2 hours in blood samples are abnormal. The renal threshold for glucose is 180 to 200 mg/100 mL of blood, and, therefore, sugar should not appear in the urine of normal subjects in the tolerance test.

Hyperglycemia and decreased glucose tolerance are seen in diabetes mellitus (to 500 mg/100 mL) and hyperactivity of the adrenal, pituitary and thyroid glands. *Hypoglycemia*, with a blood-sugar value of <60 mg/100 mL and increased glucose tolerance, is encountered in insulin overdose, glucagon deficiencies and hypoactivity of various endocrine glands. Intravenous glucose tolerance studies are used to circumvent defective absorption of glucose in the gastrointestinal tract, eg, in steatorrhea.

Monitoring hemoglobin A_{1c} is another way to follow patients with hyperglycemia. This is more specific for diagnosing diabetes but less sensitive than the glucose tolerance test.[15] Normally, hemoglobin A_{1c} accounts for 3 to 6% of the total hemoglobin while in diabetics it is 6 to 12%. The concentration of Hgb A_{1c} in the blood reflects the patient's carbohydrate status over a period of time, providing a marker for hyperglycemia. *Pancreatic function tests* include studies on IV and oral glucose, glucagon and tolbutamide tolerance. The beta cells of pancreatic islet tissue secrete insulin and the alpha cells secrete glucagon, a substance antagonistic to insulin and having a hyperglycemic effect induced by its glycogenolytic action. In *glucagon tolerance studies* the effect of parenteral administration of glucagon on blood-sugar values is useful in the diagnosis of pancreatic and hepatic function. *Insulin and tolbutamide tolerance studies* are used in the diagnosis of endocrine disorders, differentiation of insulin-resistant diabetics and determination of functional hypoglycemia and islet-cell tumors.

Galactosemia, the presence of galactose (>4.5 mg%) in blood, is usually due to an inborn error of galactose metabolism. Congenital deficiencies in galactokinase or galactose 1-phosphate uridyl transferase result in inadequate galactose metabolism with accumulation of galactose 1-phosphate in the liver. Oral administration of galactose in galactosemia leads to a decrease in blood glucose and an increase in concentrations of galactose in the urine and blood. Galactose is measured by estimation of NADH liberated in the conversion of galactose to galactonolactone in the presence of NAD and galactose dehydrogenase. Deficiencies in intestinal disaccharidases such as lactase will preclude efficient conversion of lactose to galactose and glucose, and oral administration of lactose will cause no increase in blood galactose and usually produce diarrhea. Galactose-loading studies are useful in the diagnosis of toxic or inflammatory conditions of the liver. In hepatic cirrhosis, there is a decrease in the galactose-metabolizing capacity of the liver due to the inhibition of hepatic diphosphogalactose-4-epimerase.

Lactic acid is a product of glucose metabolism; it is converted into pyruvic acid and NADH by lactic dehydrogenase (LDH) in the presence of NAD. Blood lactic acid is estimated by reaction with LDH to form pyruvate and NADH; the NADH level is determined spectrophotometrically at 340 nm and is a function of lactic acid concentration. It is elevated (>20 mg/100 mL) following exercise, anesthesia and certain types of acidosis. The *blood lactate/pyruvate* ratio should be calculated in order to determine the presence of excess lactic acid in the blood in acidosis, thiamine deficiency and decompensated heart disease.

Blood pyruvic acid is determined by the reverse procedure; ie, the conversion of pyruvate to lactate in the presence of LDH and NADH. Normal blood pyruvic acid ranges from 0.6 to 1.3 mg/100 mL by chemical methods and 0.3 to 0.7 mg/100 mL by enzymic procedures.

Nonprotein Nitrogen (NPN) Compounds—These refer to all nitrogen-containing compounds in biological fluids exclusive of protein, including nitrogen from amino acids, low-molecular-weight peptides, urea, nucleotides, uric acid, creatinine, creatine and ammonia. Blood NPN usually is determined by digesting a protein-free blood filtrate with sulfuric acid in the presence of a catalyst (SeO_2) to convert nitrogen to ammonium sulfate (Kjeldahl digestion—see page 444); the excess acid is neutralized and ammonia determined by Nesslerization or reaction with alkaline hypochlorite.

The normal blood NPN is 25 to 45 mg/100 mL (48% urea N, 14% amino acid N, 4% creatine N, 1% creatinine N, 3% uric acid N and 30% residual N). In renal damage, NPN is elevated to values ranging from 60 to 500 mg/100 mL (*azotemia*). As variations in NPN mainly reflect alterations in blood urea nitrogen (BUN), urea determinations are more sensitive and preferred as a guide to kidney function.

The primary pathway of nitrogen metabolism in man is the synthesis of urea from ammonia in the liver and then rapid renal excretion of urea. In renal disease (*nephritis*), the excretion of urea is diminished and blood NPN and BUN are increased. In BUN procedures, *urea* is converted enzymatically to ammonia by urease; the ammonia then is determined by Nesslerization, reaction with phenol-alkaline hypochlorite, aeration into standard acid and subsequent titration or reaction with salicylate-nitroprusside reagent at pH 12 in the presence of alkaline dichloroisocyanurate to form a green chromogen which can be estimated colorimetrically. The ammonia also can be estimated by spectrophotometric determination of NAD produced in the conversion of ammonia and α-ketoglutarate to glutamate by NADH-L-glutamate dehydrogenase. Direct chemical determinations of urea are based on the reaction with 2,3-butanedione in an acid medium (Fearon reaction).

BUN (normal ≈ 5 to 25 mg/100 mL) is increased in chronic and acute nephritis, metallic poisoning and cardiac failure; reduced levels occur in rapid dehydration or following diuresis. In severe liver damage due to diminished urea formation, an increase in blood ammonia and decrease in BUN are observed. Urine urea output (6 to 17 g/day) is an index of *glomerular filtration rate* (*GFR*) and kidney function. Increased dietary protein and gastrointestinal hemorrhage will increase urine urea. Decreases in urea excretion involve either tubular reabsorption or secretion defects.

The *nitrogen balance* represents the balance between nitrogen input or produced (N_{in}) and nitrogen excreted (N_{out}); in normal individuals $N_{in} \approx N_{out}$. N_{out} is regulated by renal GFR; in renal disease GFR is decreased, $N_{in} > N_{out}$ and BUN is increased. The rate of urinary excretion of parenterally administered dyes (phenolsulfonphthalein), inulin sodium, *p*-aminohippurate and mannitol are sensitive indices of GFR in *renal clearance studies*.

Creatine (methylguanidoacetic acid) and *creatinine* (creatine anhydride) are involved in the physiology of muscle

contraction. Creatine phosphate is an intracellular source of high-energy phosphate bonds via the reaction of ATP and creatine kinase. Creatinine is the waste product of creatine metabolism and is the normally excreted compound.

Serum creatinine is determined by reaction with alkaline picrate to form a red chromogen. These values usually represent 20 to 30% of noncreatinine-interfering substances. Absolute determinations can be made by the absorption of creatinine from protein-free blood filtrates on aluminum silicate prior to the final determination. Drugs causing nephrotoxicity result in a slight increase in creatinine and those which interfere with color formation in the reaction include bromosulfophthalein (BSP), phenolsulfonphthalein (PSP), acetoacetate, ascorbic acid, levodopa, methyldopa, glucose and fructose. Creatine is determined after hydrolytic conversion to creatinine with boiling, aqueous picric or hydrochloric acid.

Renal clearance of endogenous creatinine is related to GFR and is normally 1 to 2 g/day (creatinine coefficient = 20 to 26 mg/kg/24 hr). Normal serum creatinine is 1 to 2 mg/100 mL; creatine 0.2 to 1.0 mg/100 mL. Higher values (5 mg/100 mL) indicate glomerular damage or cardiac insufficiency.

Uric acid is a catabolite of purine metabolism as derived from nucleic acids or nucleotide cofactors. Direct methods for determining uric acid involve the reaction with alkaline phosphotungstic acid to form a "tungsten blue," which is estimated colorimetrically. In another method, alcoholic NaOH is added to a protein-free filtrate to eliminate interfering reducing substances (ascorbic acid, glutathione) prior to the reduction of uric acid with acid copper chelate to form a cupric chromogen complex.

In indirect procedures, uric acid is hydrolyzed by the enzyme uricase; the decrease in absorbance at 290 to 293 nm is a function of the initial concentrations of uric acid. The normal blood value is 1.5 to 6.0 mg/100 mL. It is elevated in renal disease, gout due to increased metabolic pools of uric acid and leukemia as a result of increased turnover of cellular nucleoprotein.

Amino acid determinations in blood are performed by conventional colorimetric ninhydrin techniques or reaction with alkaline β-naphthoquinone-4-sulfonate. Normal plasma values range from 3.9 to 7.8 mg/100 mL. A variety of metabolic disorders may be detected by analyzing for increased levels of specific amino acids in the urine or blood. Total urine amino acids are determined by formol titration; formaldehyde reacts with basic amino groups and thus permits subsequent titration of the acidic groups of the amino acids. Daily excretion of amino acid nitrogen ranges from 100 to 400 mg, constituting 1 to 2% of total urine nitrogen.

The identification and quantitation of specific amino acids in the blood and urine are accomplished by paper, thin-layer (TLC), column and ion-exchange chromatographic and electrophoretic separation of electrolytically desalted blood or urine samples. See Chapter 29.

Abnormal amino acid metabolism (*aminoacidopathies*) usually results in the presence of abnormal quantities of specific amino acids in the urine (aminoaciduria).

The aminoacidurias are divided into two main groups:

1. *Primary overflow aminoaciduria* in which blood amino acids are elevated [phenylketonuria (PKU), maple syrup urine disease (MSUD), tyrosinosis and alkaptonuria].
2. Aminoacidurias characterized by elevated amino acid urine levels with normal blood levels (*transport diseases* with a defect in the kidney tubule—eg, cystinuria—and "no-threshold" aminoaciduria in which the kidney has no mechanism for reabsorbing the amino acid involved—eg, homocystinuria).

PKU, a disease characterized by mental deficiency, is associated with the presence of phenylpyruvic acid in the urine and elevated serum phenylalanine levels due to a hereditary (autosomal recessive) deficiency of hepatic phenylalanine hydroxylase, which converts phenylalanine to tyrosine. The availability of treatment through dietary intake is predicated upon early detection. Many states have passed legislation for mass-screening for PKU in all infants. The Guthrie test is performed by placing filter paper discs impregnated with serum or blood on the surface of an agar culture medium containing β-(2-thienyl)alanine at a concentration sufficient to inhibit the growth of *B subtilis*. *Phenylalanine* will reverse this inhibition and the Bacterial Inhibition Assay (BIA) is a direct measure of this amino acid. Serum phenylalanine determinations also can be performed by estimating the fluorescence of a complex with ninhydrin and copper in the presence of L-leucyl-L-alanine.

MSUD is characterized by the odor of the urine and rapidly is fatal to infants. It is associated with a deficiency in the oxidative decarboxylation of α-keto acids leading to an accumulation of both the keto and amino acids in the blood and urine (valine, leucine, isoleucine). TLC and BIA assays can be used to detect MSUD.

Alkaptonuria is a rare, hereditary disease in which homogentisic acid cannot be metabolized further due to a lack of homogentisic acid oxidase. This causes homogentisic aciduria, ochronosis and arthritis.

In *Hartnup disease*, indole and tryptophane appear in the urine due to defective renal and intestinal absorption of tryptophane. Tryptophane is an intermediary metabolite in the synthesis of serotonin (5-hydroxytryptamine) and 5-hydroxyindole acetic acid (HIAA). Excessive production of *serotonin* and the presence of its *HIAA metabolite* in the urine are associated with metastatic carcinoid tumors. HIAA is measured after removal of interfering keto acids with dinitrophenylhydrazine, extraction and estimation with nitrosonaphthol reagent.

Routine screening tests for congenital metabolic defects and the substance under test in the newborn include PKU (phenylalanine), MSUD (leucine), tyrosinemia (tyrosine), homocystinuria (methionine), histidemia (histidine), valinemia (valine), galactosemia (galactose or galactose uridyltransferase), orotic aciduria (orotidine-1-phosphate decarboxylase), arginosuccinuria (arginosuccinic lyase), hereditary angioneurotic edema (C^1-1-esterase inhibitor) and sickle-cell disease (hemoglobin S).

The analyses for these substances are based on BIA, metabolite bacterial inhibition assay (MIA), enzyme auxotraph bacterial assay (ENZ-Aux), fluorescent spot tests or TLC and electrophoresis.

Proteins—The *plasma proteins* (albumins, globulins and fibrinogen) are involved in nutrition, electrolyte and acid-base balance, transport mechanisms, coagulation, immunity and enzymatic action. *Total plasma proteins* may be determined by Kjeldahl, Nessalerization, specific ion-pair (bromcresol green dye plus albumin) or biuret procedures. The last technique is based on the reaction of —CONH— groups joined by carbon or nitrogen linkages in protein with alkaline copper sulfate to yield the biuret complex which can be estimated colorimetrically. Total protein also can be estimated by specific gravity, refractometric or UV spectrometric methods. These methods are subject to large errors in the presence of a pathology involving increased glucose, lipid, urea or abnormal protein concentrations.

The *albumin-globulin (A/G) ratio* is determined by the biuret method after precipitation of the globulins with a sodium sulfate-sulfite reagent. The normal range is 5.5 to 8.0 g% total protein with an A/G ratio of 1.4 to 2.4. Changes in total protein and A/G ratio occur in kidney and liver disease, hemorrhage, dehydration, rheumatoid arthritis and multiple myeloma. Gastrointestinal albumin loss, as seen in GI bleeding, ulcerative colitis, sprue and enteritis, can be

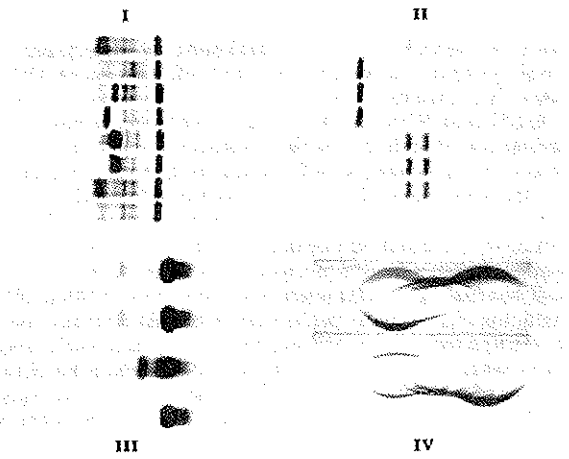


Fig 28-4. Electrophoretic separation of serum proteins (I), isoenzymes (II), hemoglobins (III), and immunoelectrophoresis of plasma protein (IV) (courtesy, Spinco).

detected by monitoring fecal radioactivity after IV injection of ^{51}Cr-human serum albumin.

The physiochemical properties of the plasma proteins—mol wt (68,000 to 300,000) and isoelectric point (pH of minimum solubility and ionic neutrality)—provide the basis for the electrophoretic separation of plasma proteins (Fig 28-4). The plasma sample is spotted on a paper or cellulose acetate strip, or in a polyacrylamide gel (disc or gel electrophoresis) at pH 8.6.

At this pH the proteins are electroanionic and, under the influence of electric current, will migrate to the anode at a rate dependent on their isoelectric point and, in the case of cellulose acetate or gel electrophoresis, their molecular size. The strips are then stained with a protein dye (bromophenol blue, Amidoschwarz or Ponceau S), and the concentrations of the various proteins are estimated by densiometric scanning.

The normal ranges for the major proteins are (in g%): albumin 3.8 to 5.0; total globulin, 2.0 to 3.9; α_1-globulin, 0.1 to 0.5; α_2-globulin, 0.5 to 0.9; β-globulin; 0.5 to 1.2; γ-globulin, 0.7 to 1.6.

Ordinary electrophoresis does not identify the subgroups of *immunoglobulins*, IgA, IgM, IgG and IgE. This is accomplished by immunoelectrophoresis, a process involving electrophoresis and immunodiffusion. The sample is electrophorized in an agar gel (zone electrophoresis) and then antiserum to the specific Ig or to total globulins is placed in a trough aligned parallel to the axis of the original electrophoresis. The serum proteins and antisera diffuse toward each other and form precipitin (antigen-antibody complex) lines. Ordinary cellulose acetate or gel electrophoresis will permit the recognition of diffuse, polyclonal elevation of serum immunoglobulins seen in chronic infections, isolated M-protein peaks of macroglobulinemia and multiple myeloma and absent gamma component in a hypogammaglobulinemia or agammaglobulinemia. Immunoelectrophoresis will indicate specific Ig abnormalities or, by noting the presence of any displacement, bowing or broadening of the precipitin band will aid in the diagnosis of the paraimmunoglobulin monoclonal diseases such as multiple myeloma, macroglobulinemia or chronic lymphatic leukemia.

Radial immunodiffusion is a simple process which also can be used for quantitation of IgA, IgM and IgG.[16] It is performed by incorporating the antibody in an agar gel and then introducing the antigen or test sera into wells punched in the agar. The antigen diffuses radially out of the well into the surrounding gel media, and a visible precipitin line forms where the antigen and antibody have reacted. Quantitation of IgA, IgM and IgG aids in the diagnosis and differentiation of collagen diseases, chronic infections and liver disease. IgE is best quantitated by immunoelectrophoresis or RIA (see section on *Immunology* for the basis and principles of RIA).

Nephelometric techniques detect immunological constituents by measuring the light-scattering properties of various antigen-antibody complexes in a test solution. The Hyland system measures the amount of laser-beam deflection at an angle by employing a photomultiplier tube which is sensitive in the red region of the spectrum. Results are calculated by an electronic-screening system and read in percent relative light-scatter on a digital readout.

Automated electrophoresis equipment offers computer-controlled sample application, staining options, densitometry and pattern interpretation for serum proteins and isoenzymes.

Enzymes—Enzymes are proteins whose biological function is the catalysis of chemical reactions in living systems. Enzymes combine with the substances on which they act (substrates) to form an intermediate enzyme-substrate complex which is then converted to a reaction product and liberated enzyme, which continues its catalytic function. Enzymes are highly specific; a few exhibit absolute specificity and catalyze only one particular reaction, while others are specific for a particular type of chemical bond, functional group or stereoisomeric structure.

Most serum enzymes of clinical significance are intracellular in origin and are elevated in hyperactivity disease, malignancy or injury to cardiac, hepatic, pancreatic, muscle, bone and tissue. As the specific tissue involved will determine the type of enzyme that will be elevated, such determinations are valuable diagnostic tools in the differentiation of various pathological states.

Enzymes are named and classified according to the type of reaction that they catalyze, and to their substrate specificities. Enzyme activity usually is expressed in International Units (IU) where 1 unit (U) is that amount of the enzyme which will catalyze the transformation of 1 μmole of substrate/min at definite temperature, pH and substrate-concentration conditions. Refer to Chapter 52 for a more complete discussion of enzymes.

Transferases are enzymes that catalyze the transfer of amino or phosphate groups from one compound to another. Aspartate aminotransferase (AST) and alanine aminotransferase (ALT) are important in clinical diagnosis. These enzymes catalyze the transfer of the amino group from glutamic acid to keto acids (oxaloacetic or pyruvic) to form aspartic and α-ketoglutaric acids with AST (aspartate aminotransferase) and alanine and α-ketoglutaric acid with ALT (alanine aminotransferase).

Colorimetric methods are based on an estimation of the reaction products (oxaloacetic or pyruvic acid) with dinitrophenylhydrazine, or substrate (α-ketoglutaric acid) by coupling with 6-benzamido-4-methoxy-*m*-toluidinediazonium chloride.

Spectrometric methods are based on the reaction of the product pyruvate with lactic dehydrogenase and NADH, or of oxaloacetate with malic dehydrogenase and NADH. The rate of NADH utilization is measured by the decrease in absorbance at 340 or 360 nm and is directly proportional to transaminase activity.

Normal AST and ALT levels are <40 mU/mL. AST is present in large amounts in liver, cardiac and skeletal muscle, whereas ALT is found primarily in liver tissue. AST is elevated in myocardial infarction and Duchenne muscular dystrophy; AST and ALT are increased in liver disease, acute toxic or viral hepatitis, infectious mononucleosis, obstructive jaundice and hepatic cirrhosis.

Creatine kinase (*CK*) is a transferase found in muscle and brain tissue. It catalyzes the transfer of phosphate groups from creatine phosphate to adenosine diphosphate (ADP) to form adenosine triphosphate (ATP). Activated CK activity is measured by following the increase of ATP in the creatinine phosphate-ADP reaction in the presence of glutathione or cysteine thiol activators. The ATP can be measured by the fluorimetric determination of light emitted by luciferinase conversion of luciferin to adenyl-oxyluciferin in the presence of ATP. Normal serum levels are <50 mU/mL; it is elevated in myocardial infarction and Duchenne muscular dystrophy, but remains at normal levels in liver disease.

Ornithine transcarbamylase (*OTC*) in serum is the only enzyme of the urea cycle which has been used in the clinical investigation of liver disease. It catalyzes the conversion of ornithine to citrulline. The normal serum value is 0 to 0.4 mU/mL.

Oxidoreductases or *dehydrogenases* are enzymes that catalyze hydrogen transfer in cellular oxidation processes. *Lactic* (*LDH*), *α-hydroxybutyric* (*HBDH*), *malic* (*MDH*), *glutamic* (*GLDH*), *isocitric* (*ICDH*) *and sorbitol* (*SDH*) *dehydrogenases* are of diagnostic importance in myocardial and liver disease.

LDH catalyzes the reversible conversion of pyruvic to lactic acid in the presence of NADH. The activity may be estimated colorimetrically by forming the pyruvic acid hydrazone with 2,4-dinitrophenylhydrazine; spectrometric or fluorimetric estimation of NADH in this reaction also is used to estimate enzyme activity. The normal serum LDH value is <200 mU/mL (pyruvate → lactate) and <50 mU/mL (lactate → pyruvate). LDH is increased to a much greater extent and for a more prolonged period than AST or CK in myocardial infarction; it also is increased to varying degrees in certain types of hepatic disease, disseminated malignancies, pernicious anemia and muscular dystrophy.

Recent advances in protein chemistry and technical methodology have led to fractionation of enzymes, previously thought to be homogeneous, into heterogeneous moieties. These multiple-molecular forms of enzymes (*isoenzymes*) have similar substrate specificity but different biophysical properties. LDH, MDH, CK, phosphatases and leucine aminopeptidase exist in isoenzyme forms.

CK isoenzymes are important in the early detection of myocardial damage. Two CK molecular subunits, M and B, produce three isoenzymes: CK-MM found primarily in skeletal muscles, CK-MB in the myocardium and CK-BB primarily from the brain. After acute myocardial infarction (MI), CK-MB appears in the serum in approximately 4 to 6 hours, reaches peak activity at 18 to 24 hours and may disappear within 72 hours. Diagnostic testing of MI includes CK and LDH isoenzymes. Early detection of CK-MB allows the management of myocardial infarcts with agents such as streptokinase or tissue plasminogen activator (TPA). The methods of assessment include electrophoresis, column chromatography and immunoinhibition.

Serum contains five LDH isoenzymes, each a tetramer composed of one or two monomers. LDH 1 and 2 are found in preponderance in heart, kidney and RBC; whereas liver and skeletal muscle largely contain LDH 4 and 5. Intermediate forms prevail in lymphatic tissues and many malignancies. The fractionation of LDH isoenzymes is important in the differential diagnosis of cardiac, muscle and liver disease. It can be accomplished with DEAE-cellulose chromatography, electrophoresis, sulfite or urea inhibition of specific isoenzymes, thermal stability and substrate-concentration requirements.

HBDH reduces α-ketobutyric acid to α-hydroxybutyric acid in the presence of NADH; estimation of the α-keto acid via hydrazone formation or NADH is the basis of activity measurements. The normal serum HBD level is <140 mU/mL; it is elevated in myocardial infarction. LDH 1 is high in HBDH activity. The ratio of total LDH/HBDH often is used in place of LDH isoenzyme determination. Ratios >0.8 are seen in myocardial infarction and <0.6 in acute liver damage.

MDH and *SDH*, in the presence of NAD, catalyze the conversion of malate or sorbitol to oxaloacetate or fructose, respectively. They are of diagnostic value in MI (MDH >48 mU/mL) and acute liver injury (SDH >96 mU/mL).

ICDH oxidizes isocitrate, in the presence of NADP or NAD, to α-ketoglutarate; it is elevated (>5.0 mU/mL) in acute hepatitis.

Hydrolases are enzymes that catalyze the addition of the elements of water across the bond which is cleaved. *Amylases, lipases, phosphatases, 5′-nucleotidase, γ-glutamyl-transpeptidase* and *leucine aminopeptidase* are specific examples of clinically important hydrolases.

Salivary and pancreatic *amylases* hydrolyze the substrate starch to maltose and dextrins. Amylase activity can be measured by procedures based on the loss in certain properties of starch as it is hydrolyzed (*amyloclastic*), or by the generation of reducing substances (*saccharogenic*). The amyloclastic methods use the decrease in viscosity and turbidity of hydrolyzed water-soluble starch substrates, or the reaction of starch with iodine as the method of estimation. A newer procedure uses the colorimetric determination of water-soluble dye-dextrin fragments released by amylolytic hydrolysis of a cross-linked, water-insoluble, dye-starch polymer. The saccharogenic methods determine the reaction products (reducing sugars) by a previously described methodology. The normal serum level is 140 mU/mL; elevations are noted in acute pancreatitis, acute abdominal conditions (perforated peptic ulcer, common bile-duct obstruction) and salivary gland disease.

Lipases catalyze the conversion of triglycerides to glycerol and fatty acids. Clinical determinations are based on the titrimetric analysis of fatty acids liberated from an emulsified olive oil substrate, or fluorimetric estimation of fluorescein liberated from a fluorescein fatty acid ester substrate. Serum lipase is increased in pancreatic carcinoma.

Phosphatases catalyze the hydrolysis of orthophosphoric acid esters, and are classified according to the pH of optimal activity into alkaline or acid phosphatases. Activity (alkaline, pH 8 to 10; acid, pH 4 to 6) is measured with phenyl phosphate, glycerophosphate, *p*-nitrophenyl phosphate or thymolphthalein monophosphate substrates. With the latter two chromogenic substrates, the amount of *p*-nitrophenol or thymolphthalein liberated by phosphatase hydrolysis is estimated colorimetrically in an alkaline medium. With a glycerophosphate or phenyl phosphate substrate, the liberated phosphorus is determined by molybdenum blue formation with phosphomolybdic-phosphotungstic acids; phenol also may be estimated with 4-aminoantipyrine or Folin-Ciocalteau reagent.

Acid phosphatase activity may be differentiated by the use of inhibitors in the assay mixture; formaldehyde has no effect on acid phosphatase of prostatic origin, but it inhibits other acid phosphatases, while tartrate is a selective inhibitor of the prostatic enzyme. *Acid phosphatase* is of a primary diagnostic value in metastatic carcinoma of the prostate.

Normal values for *alkaline phosphatase* activity depend on the substrate used; elevations in osteomalacia and in bone tumors depend on the degree of osteolytic or osteoblastic activity. The enzyme (isoenzyme) also is elevated in obstructive jaundice, bone and liver disease.

The enzyme *5′-nucleotidase* is an alkaline phosphomonoesterase that hydrolyzes nucleotides with a phosphate radical attached to the 5′-position of the pentose (eg, adenosine monophosphate). The normal serum value is 17 mU/mL; it is elevated in hepatic disease.

Leucine aminopeptidase (LAP) is an exopeptidase which

hydrolyzes the peptide bond adjacent to a free amino group. It liberates amino acids from the *N*-terminal group of proteins and polypeptides in which the free amino group is a L-leucine residue. Activity is determined by spectrophotometric estimation following hydrolysis of the amide bond of a leucinamide substrate at 238 nm. Clinical estimations usually are performed on synthetic substrates, and since there is no correlation between cleavage of leucinamide and these substrates, the LAP-like activity is designated *leucine arylamidase*. A fluorometric determination of naphthylamine liberated from a leucyl-β-naphthylamide substrate or colorimetric determination of *p*-nitroaniline liberated from leucino-*p*-nitroanilide substrate also has been used. The normal value is 8 to 22 mU/mL; it is elevated in the last trimester of pregnancy, hepato-biliary disease and pancreatic carcinoma.

Serum *γ-glutamyl transpeptidase* (γGT) is increased in diseases of the liver, bile ducts and pancreas. Together with alkaline phosphatase, LAP and 5′-nucleotidase, γGT usually is tested in the group of cholestasis-indicating enzymes. The assay is based on the hydrolysis of γ-glutamyl-*p*-nitroanilide.

Serum lysozyme (muramidase) activity is increased in certain types of leukemia. Serum arginase, an enzyme which hydrolyzes arginine to ornithine and urea, and serum guanase are sensitive indicators of hepatic necrosis.

Lyases are enzymes which split C—C bonds without group transfer. *Aldolase* is a glycolic lyase which catalyzes the reversible splitting of fructose 1,6-diphosphate to form dihydroxyacetone phosphate and glyceraldehyde 3-phosphate. In the estimation of activity, the triose phosphate reaction products are hydrolyzed with alkali and the resultant trioses are reacted with 2,4-dinitrophenylhydrazine to form chromogenic hydrazones for colorimetric analysis. A spectrophotometric estimation is made by coupling the aldolase reaction products with a dehydrogenase acting on one of the triose phosphates and measuring concomitant changes in NADH. The normal value is <8 mU/mL; it is elevated in muscular dystrophy, polymyositis and acute hepatitis.

The significance of serum-enzyme changes in hepatitis is seen in Fig 28-5 and enzyme activity following myocardial infarction in Fig 28-6.

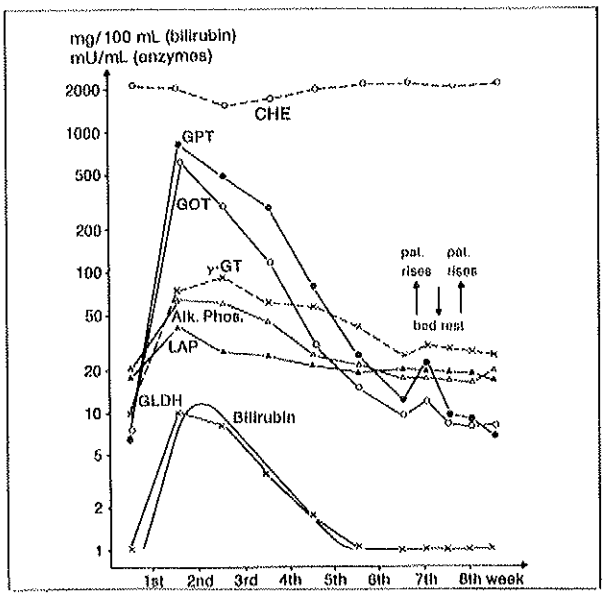


Fig 28-5. Typical course of alterations in serum enzyme activity in acute viral hepatitis (courtesy, Schmidt E, Schmidt FW *Med Welt 21:* 805, 1970).

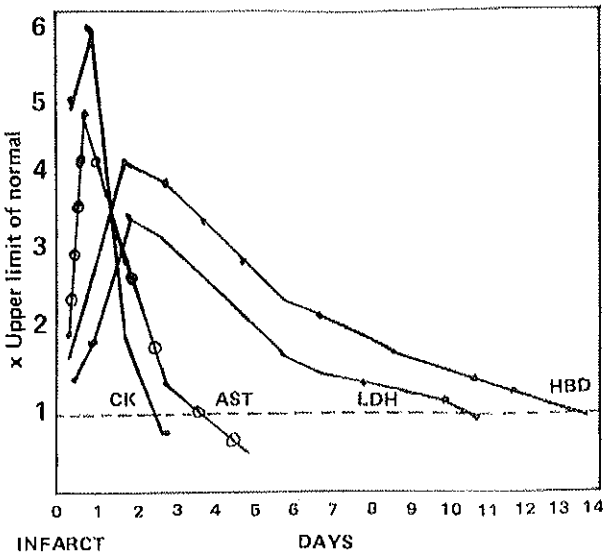


Fig 28-6. Serum enzymes following myocardial infarction, AST, CK, LDH and HBD are compared.

Lipids—The major classes of blood lipids are *fatty acids, cholesterol, triglycerides, phospholipids and lipoproteins.* Hyperlipidemia is not a single aberration and there are a number of different hyperlipidemic states. Lipid-profile tests include measurements of cholesterol, triglyceride, phospholipids and determination of lipoprotein phenotypes.

Cholesterol, a sterol molecule, is an essential substance in steroid-hormone synthesis by the adrenal cortex and bile acid production in the liver. It exists in blood as the free sterol and as cholesterol esters of fatty acids.

In the determination of *total cholesterol*, the serum is extracted with an alcohol-ether mixture and the cholesterol estimated colorimetrically after reaction with acetic anhydride-sulfuric acid reagent (Liebermann-Burchard reaction). The precipitation of free cholesterol with digitonin will differentiate free from esterified cholesterol. Chromatographic separation of cholesterol from its esters on alumina, silicic acid or magnesium silicate columns with organic solvents also has been used.

Gas chromatographic procedures have resulted in the separation and quantitation of cholesterol, its metabolites and precursors; this is a type of partition chromatography in which a volatilized sample is partitioned between a liquid stationary phase and a mobile gas phase. The normal-adult total-serum-cholesterol level is 150 to 270 mg/100 mL; it is increased in hyperlipemia and specifically in hyper-β-lipoproteinemia, nephrosis, diabetes mellitus and myxedema, and decreased in hyperthyroidism and hepatic disease. Free cholesterol comprises 20 to 40% and the ester fraction 60 to 80% of the total serum cholesterol.

Phospholipids are "compound" or "heterolipids" which contain phosphorus, a nitrogen base and a long-chain fatty acid. Lecithin (phosphatidylcholines) and cephalin (phosphatidylethanolamine or serine) are the principal plasma phospholipids, which normally comprise one-third of the total plasma lipids. They usually are bound to lipoproteins. These serum lipids are extracted into an alcohol-ether mixture, digested with sulfuric acid-hydrogen peroxide and the liberated phosphorus determined by colorimetric techniques. The normal lipid phosphorus is 6 to 11 mg/100 mL; about one-half is lecithin. The average ratio of cholesterol to lipid phosphorus when cholesterol is normal is 21. Phospholipid changes usually are associated with cholesterol changes and are of interest in coronary artery and liver diseases and the hyperlipoproteinemias.

Sphingolipids differ from lecithin and cephalin. They are phosphate esters of sphingosine bound to choline or ethanolamine and primarily are found in brain tissue (eg, sphingomyelin, galactolipin). The *ratio of lecithin to sphingomyelin* (L/S) in amniotic fluid or resuscitated amniotic fluid from the oral cavity of the newborn is an accurate assessment of fetal maturity and the respiratory-distress syndrome. Changes in phospholipid biosynthesis during gestation reflect the aging of the fetal lung, as the L/S ratio normally increases.

Tay-Sachs disease is a lipid-storage disease in which the central nervous system degenerates because of the progressive intraneuronal accumulation of excess amounts of the sphingolipid ganglioside GM_2. The accumulation of GM_2 in Tay-Sachs disease has been shown to be caused by a lack of the enzyme hexosaminidase A. Therefore, the measurement of serum, WBC or amniotic fluid *hexosaminidase A* is important in evaluating carriers and in diagnosing Tay-Sachs disease in the fetus.

Both hexosaminidase A (heat-labile) and hexosaminidase B (heat-stable) can catalyze the conversion of 4-methylumbelliferyl-*N*-acetylgalactosamine (a synthetic substrate) to *N*-acetylgalactosamine and 4-methylumbelliferone. The cleavage product, 4-methylumbelliferone, fluoresces under ultraviolet radiation and the intensity of the fluorescence is a measure of the activity of the enzyme. In noncarriers, 50 to 75% of the total hexosaminidase activity is heat-labile (hexosaminidase A), and in carriers 20 to 45% of the total hexosaminidase activity is heat-labile.

The blood fatty acids occur in esterified (EFA) and nonesterified (NEFA) forms. *Triglyceride* determinations are of value in differentiating the hyperlipidemic states, ie, essential (diet-induced) hypertriglyceridemia from familial hypocholesterolemia with or without triglyceridemia. After the preliminary separation from phospholipids, triglycerides most often are determined in terms of their glycerol moiety. The glycerol released by saponification is oxidized to formaldehyde and the latter determined by fluorimetric or colorimetric procedures. Triglycerides also can be determined by coupling the glycerol liberated from lipase/α-chymotrypsin treatment of serum with a glycerol kinase-pyruvate kinase-LDH system and spectrometric estimation of NADH. Normal triglyceride levels are 110 to 140 mg/100 mL. An increase in triglycerides will produce a milky appearance in serum (lipemic). EFA analyses are based also on the reaction of alkaline hydroxylamine with esters of fatty acids to form hydroxamic acids which produce a red color with ferric chloride.

Gas chromatographic procedures have been used to quantitate the various *fatty acids;* ie, palmitic, stearic, oleic, linoleic and linolenic acids. Mono-, di- and triglycerides also can be separated into classes and quantitated by column or thin-layer chromatography, and infrared spectrometry. The total fatty acids of plasma range from 200 to 450 mg/100 mL in the fasting state; they are derived from glycerides, cholesterol esters and phospholipids.

All the lipids in plasma circulate in combination with protein. The free fatty acids are bound to albumin and the lipids aggregate with other proteins to form *lipoproteins.* Electrophoresis and ultracentrifugation are the principal methods used to separate and identify lipoprotein families. *Chylomicrons* (S_f >400), *pre-β-lipoproteins* (S_f 20–400), *β-lipoproteins* (S_f 0–20) and *α-lipoproteins* are the four major classes in order of increasing density and migration on cellulose acetate electrophoresis. Chylomicrons are representative primarily of dietary or exogenous triglycerides, pre-β-lipoproteins of endogenous glycerides, β-lipoproteins of cholesterol and its esters and α-lipoproteins of cholesterol and phospholipids. Abnormal lipoproteins that may appear in plasma include floating β-lipoproteins, lipoprotein X and complexes of normal lipoproteins with IgA and IgG myeloma proteins (autoimmune hyperlipoproteinemia). Age, sex, diet, fasting, posture changes and trauma can alter the lipid profile.

The lipoprotein classes usually are separated by paper, agarose or cellulose acetate electrophoresis. The strips are stained with fat-soluble dyes (Sudan Black or Oil Red O) and quantitated by densiometric scanning. Primary hyperlipoproteinemias are classified into normal and five abnormal types based on cholesterol and triglyceride levels and lipoprotein analysis. Hyperchylomicronemia (Type I), hyper-β-lipoproteinemia (Type II), broad β-band (Type III), hyper-pre-β-lipoproteinemia (Type IV) and hyper-pre-β-lipoproteinemia and chylomicronemia (Type V) are the major classes. Carbohydrate and fat-tolerance studies, post-heparin lipase activity and clinical symptomatology also are integrated into the diagnosis of the various subclasses. The presence or predisposition to coronary artery disease and other disease states is associated with the various types.[17]

Steroids and Other Hormones—The steroids possess a common structure, the perhydrocyclopentanophenanthrene nucleus, and include cholesterol, bile acids, androgens and the adrenocortical, adrenomedullary, estrogenic and progestational hormones.

Androsterone, dehydroepiandrosterone, etiocholan-3α-ol-17-one, 11-ketoandrosterone, 11-ketoetiocholanolone, 11β-hydroxyandrosterone and 11β-hydroxyetiocholanolone are the principal *urinary 17-ketosteroids (17KS).* These androgenic hormones are derived from the adrenal and, in males, testicular function. The principal urinary steroid metabolites in this group of androgens are found both in the free form, and as conjugates of glucuronides, sulfates or acetates. Their determination in urine involves the acid hydrolysis of the conjugates, extraction with organic solvent, reaction with alkaline *m*-dinitrobenzene (Zimmerman reaction) and colorimetric estimation of the chromogen. The individual 17KS can be separated by TLC prior to analysis to obtain further information on the individual steroids. The normal adult urine values are: male, 9 to 24 mg/day; female, 5 to 17 mg/day. Decreased excretion is seen in hypoactive disease of the pituitary, gonads and adrenals. Increased excretion is seen in hyperplasia, cancer or tumors of the adrenals.

Testosterone is the most potent androgen in blood. The measurement of urinary or serum testosterone is useful in distinguishing normal and hypogonadal males and in treating hirsutism in the female. This hormone is determined by gas chromatography, competitive protein-binding, isotope dilution or RIA procedures. Normal serum testosterone is 0.2 to 1.1 μg/100 mL in the male and <0.1 μg/100 mL in the female.

The natural *estrogenic hormones* are estradiol, estrone and estriol, produced in the gonads, adrenals and placenta. The relative amounts of the three estrogens rise and fall concomitantly during the menstrual cycle. Maternal, urinary total-estrogen excretion, especially estriol, is an indirect index of the integrity and viability of the fetoplacental unit. Analysis involves acid or glucuronidase-arylsulfatase hydrolysis of the conjugates, removal of urinary glucose if present, extraction and colorimetric or fluorimetric analysis. In the determination, after acid hydrolysis and ether extraction of the urine, the estrogens are methylated with dimethyl sulfate and chromatographically separated prior to reaction with phenolsulfuric acid to yield a red chromogen for colorimetric analysis. The normal estrogen output is 4 to 60 μg/24 hr in the female and up to 25 μg in the male. Estrogen deficiency can be related to ovarian failure and pituitary deficiency.

Progesterone is a progestational hormone which is secreted by the corpus luteum of the ovary and also by the adrenal

cortex. Serum progesterone determination is of value in the detection of ovulation and is a measure of the secretory activity of the placenta during pregnancy. Progesterone is determined in serum by RIA, double-isotope derivatization, gas-liquid chromatography or competitive protein-binding techniques. Normal, menstrual-cycle serum progesterone levels vary between 0 and 1.6 μg/100 mL.

Pregnanediol is the principal metabolite of progesterone. The urinary determination of pregnanediol excretion is an indirect index of progesterone levels but is subject to variation due to individual differences in hepatic metabolism of this hormone and is not representative of total endogenous progesterone production.

Adrenal cortex steroids include glucocorticoids, androgens, estrogens, progesterone and mineralocorticoids. Glucocorticoids can be determined as plasma cortisol (plasma 17-OH corticosteroids), urinary-free (unconjugated cortisol) or total-urinary *17-OH corticosteroids*. The latter are determined in urine as *17-ketogenic steroids* (17KGS). The 17KS in urine are reduced with borohydride to alcohols; the 17-OH steroids are oxidized with sodium bismuthate or periodate to 17KS and quantitated by the alkaline dinitrobenzene method. The 17-OH steroids can be quantitated directly by the phenylhydrazine-sulfuric acid reaction after hydrolysis of glucuronide conjugates and chromatographic purification. The 17-OH steroid analysis only determines compounds with the dihydroxyacetone side chain, such as tetrahydrocortisol or tetrahydrocortisone; the 17KGS analysis includes the 17-OH-corticosteroids with the dihydroxyacetone side chain and the pregnanetriol type of compound. Normal 17KGS daily urinary excretion is 5 to 23 mg in the male and 3 to 15 mg in the female. They are reduced significantly in myxedema and adrenal or anterior pituitary insufficiency. Plasma cortisol usually is measured by fluorimetric or gas-chromatographic procedures.

Aldosterone is the most active member of the mineralocorticoid group. The determination of urinary aldosterone is of value in differentiating benign essential hypertension from primary aldosteronism (Conn's syndrome), which is caused by an adrenal adenoma and is accompanied by hypertension. A double-isotope derivatization technique is used. Urinary aldosterone is acetylated with ^{3}H-acetic anhydride; aldosterone-^{14}C-diacetate standard is added early in the procedure. The $^{3}H/^{14}C$ specific activity of the final product is measured after chromatographic purification and is a direct measurement of aldosterone. The normal aldosterone levels of about 10 μg/day are elevated in Conn's disease and usually are associated with low serum potassium, sodium retention and low-concentration alkaline urine.

The anterior pituitary secretes three substances (*gonadotropins*) which regulate gonadal activity: *follicle-stimulating hormone* (*FSH*), *luteinizing hormone* or *interstitial cell hormone* (*ICSH, LH*) and *luteotropin* (*LTH*). The gonadotropins are glycoproteins. Bioassay methods can be used to determine gonadotrophic activity. After fractionation and isolation the urine extract is assayed in test animals as to the follicular growth of the ovaries in hypophysectomized animals or increase in testicular, ovarian or uterine weight in various animal models. RIA techniques have been developed for these gonadotropins and represent the most sensitive and precise measurement method.

Analysis of serum or urinary *placental lactogen* (*HPL*) and *chorionic gonadotropin* (*HCG*), a placental-derived protein hormone, is useful in the diagnosis of threatened abortion, hydatiform mole and choriocarcinoma. HCG, pregnanediol and progesterone as well as total and fractionated estrogens are useful in testing for pregnancy. HCG and HPL readily are measured by RIA and low values are seen in threatened abortion and intrauterine fetal death.

The increase in HCG in the serum or urine of the pregnant female is the basis of a routine *pregnancy test*. Test components consist of an antigen in the form of HCG latex particles and an HCG antiserum. When antiserum is mixed with urine containing a detectable level of HCG, it is neutralized and no agglutination of latex-antigen particles occur (*agglutination inhibition test*). The commercial application of the HCG assay gives laboratories a rapid, accurate pregnancy test by taking advantage of monoclonal antibody specificity and sensitivity. A monoclonal slide procedure on urine, Duoclon (*Organon Diagnostics*), uses two different monoclonal antibodies, one against HCG and one against the HCG_{β} subunit for maximum specificity. Agglutination indicates a positive test with a sensitivity level of 500 miU HCG/mL, detecting pregnancy a few days after conception.

Human growth hormone and *insulin* are proteins which are of diagnostic value in growth-rate studies and diabetes. They are best quantitated by RIA.

Epinephrine and *norepinephrine* are biologically active catecholamines derived from the adrenal medulla and sympathetic nerve endings. Catecholamines are measured in the blood and urine after fractionation on alumina or ion-exchange columns, oxidation at pH 3.5 or 6.0 and subsequent fluorimetric analysis. Urine catecholamines are increased to >350 μg/24 hr in adrenal medullary tissue tumors (pheochromocytoma). The normal plasma level is 2.1 to 6.5 μg/L with about 80% as norepinephrine.

Vanillylmandelic acid (*VMA*) is the urine metabolite of these two catecholamines. Its quantity in urine reflects the endogenous secretion of catecholamines. VMA can be determined colorimetrically, after extraction of the urine with ethyl acetate and diazotization with *p*-nitroaniline and ethanolamine in the presence of carbonate ion. VMA also can be measured spectrometrically following periodate oxidation to vanillin and solvent extraction. The normal output is 0 to 12 mg/24 hr.

Homovanillic acid (*HVA*) is not a metabolite of epinephrine or norepinephrine, but is produced from a common precursor, dopamine. Elevated HVA excretion is diagnostic in cases of neuroblastoma.

The biosynthesis of *serotonin* (5-hydroxytryptamine) and urinary excretion of its metabolite, 5-hydroxyindoleacetic acid (5-HIAA), are increased in argentaffine tumors. These have a very large capacity to metabolize tryptophane stores to serotonin. Urinary 5-HIAA increases from 1 to 7 mg/24 hr to as much as 1 g/24 hr in this type of tumor.

Bilirubin, a tetrapyrrole which is derived from senescent red-cell degradation, normally occurs in low concentration in the blood. In bile, it is present as the water-soluble conjugated acyldiglucuronide. In blood, bilirubin is tightly bound to plasma albumin. The reduction of bilirubin in the intestine yields urobilinogen which is, in turn, oxidized to a brown pigment—urobilin.

Serum bilirubin is determined by coupling with diazotized sulfanilic acid to form azobilirubin for colorimetric analysis. The *direct* or *conjugated bilirubin* test is performed in aqueous media; the *indirect* or *free bilirubin* analysis is performed in methanol or caffeine-sodium benzoate solution. Normal values in serum are: direct, 0 to 0.3 mg/100 mL; total, 0 to 1.5 mg/100 mL.

Clinical jaundice is a yellowing of the tissues associated with hyperbilirubinemia; in hemolytic disease of the newborn due to Rh and ABO incompatibilities, indirect serum bilirubin is elevated, whereas acute hepatitis results in increases in the direct type.

Electrolytes—The normal plasma electrolyte level is 154 mEq/L of cations and 154 mEq/L of anions. The osmotic effects of chloride, bicarbonate, sodium and potassium are important in the maintenance of normal muscle contraction and water distribution between cells, plasma and interstitial fluid.

Flame photometry, atomic-absorption spectrometry, neutron-activation analysis, X-ray fluorescence, ion-specific electrodes and colorimetric techniques are used in the identification and determination of cations or anions in biological fluids. Advances in technology have developed multiphase systems capable of measuring not only sodium and potassium but also chloride, carbon dioxide and calcium simultaneously.

Sodium and *potassium* serum concentrations are readily measured by flame photometry or highly sensitive and specific atomic-absorption spectrometry. The latter technique is similar to emission-flame photometry, except that it measures energy as it is absorbed by atoms rather than as it is emitted by atoms. Both techniques are based on the characteristic absorption or emission wavelengths of the cations. Ion-specific glass electrodes also are used for Na^+ and K^+ determinations, eliminating the use of a flame or combustible gas and can be performed on whole blood, plasma or serum.

Chloride levels in serum or urine are determined by titration with acid mercuric nitrate solution in the presence of s-diphenylcarbazone indicator. They also may be determined potentiometrically with a silver-silver chloride pH electrode assembly. The normal serum values are 135 to 155 mEq Na/L, 3.9 to 5.6 mEq K/L and 95 to 106 mEq Cl/L; urine levels are 150 to 197 mEq Na/day, 20 to 64 mEq K/day and 180 to 270 mEq Cl/day.

Serum sodium, potassium, chloride and *bicarbonate* determinations are useful indicators in adrenal cortical insufficiency, renal and cardiac failure, anuria, dehydration, alimentary tract diseases associated with diarrhea and vomiting and increased renal electrolyte excretion (diuretic therapy).

The determination of excess *chloride* (>50 mEq/L) in the perspiration of patients with pancreatic *cystic fibrosis* is an accurate diagnostic tool. Perspiration is stimulated by placing the patient's hand in a plastic bag for 15 to 20 min or, preferably, by an iontophoresis technique in which pilocarpine nitrate ions are transported through small areas of the skin to produce local perspiration. The chloride content may be quantitated with silver nitrate–potassium chromate-impregnated papers or with ion-selective electrodes.

Bicarbonate, phosphates, sodium, potassium and chloride concentrations are related to maintenance of acid-base balance in the body. The pH of the blood reflects the state of the acid-base balance and is related mathematically to HCO_3^- concentration and partial pressure of CO_2 (pCO_2) in blood by the Henderson-Hasselbach equation.

$$pH = 6.1 + \log \frac{[HCO_3^-]}{[H_2CO_3]} \qquad (2)$$

Blood pH, as measured electrometrically, has a normal range of 7.36 to 7.40 for venous samples and 7.38 to 7.42 for arterial samples. The *pCO_2 level* in blood is determined by measuring the pH of the blood at three different pCO_2 concentrations—one native to the blood and the other two obtained by equilibration with gas mixtures of known pCO_2. Blood bicarbonate levels also may be determined by measuring the amount of acid neutralized by plasma or serum and pCO_2 calculated by Eq 2. The relationship between pCO_2 and carbonic acid concentration is

$$[H_2CO_3] = 0.03 \times pCO_2$$
$$\text{mM per L} \qquad \text{torr} \qquad (3)$$

The role of oxygen and hemoglobin in respiration has been discussed previously. Measurements of blood pH and CO_2 content are used in differentiating respiratory acidosis (low pH, high CO_2) from metabolic acidosis (low pH, low CO_2).

Blood oxygen (*pO_2*) and *percent oxygen saturation* are measured by a polarographic method; the blood sample is placed in a chamber and separated from a combined platinum and silver-silver chloride electrode by a polypropylene membrane. By diffusion through the membrane, equilibrium is established between the pO_2 of the blood and a film of solution in contact with the electrode. A current, which is proportional to blood pO_2, is generated after the application of a polarizing voltage.

Calcium and phosphorus are important minerals in the processes of bone calcification, nerve irritability, muscle contraction and blood coagulation. Calcium is present in plasma as an ultrafilterable (ionic and nonionic) form and a protein-bound fraction. Blood phosphorus consists of inorganic phosphorus, organic phosphate ester (G6P, ATP) and phospholipids.

Serum and urine calcium levels are determined routinely by titration with EDTA or EGTA using a fluorescent calcein or calcichrome indicator. Other methods are based on the colorimetric analysis of calcium-methylthymol blue complex in the presence of 8-quinolinol to prevent interference by magnesium. Bis-(*o*-hydroxyphenylimino)ethane forms a colored complex with calcium and, in the presence of polyvinylpyrrolidone to inhibit phosphate interference, is a sensitive and specific method for calcium. Calcium is determined best by atomic-absorption spectrometry. As with all cations, calcium can be determined by emission- or absorption-flame photometry or ion-selective electrodes.

Inorganic phosphorus levels are determined by reaction with acid molybdate reagent to form phosphomolybdic acid which, in turn, is reduced with aminonaphtholsulfonic acid or *p*-dimethylaminophenol sulfate to give a blue complex which is estimated colorimetrically. Normal serum levels are 2.5 to 4.5 mg P/100 mL and 9 to 11 mg Ca/100 mL.

Calcium levels are decreased and phosphorus increased in hypoparathyroidism; an opposite effect is seen in hyperactivity of this gland. In rickets and osteomalacia, the concentrations of both elements are decreased. In establishing primary hyperparathyroidism and other causes of hypercalcemia, daily measurements for ionized calcium (Ca^{2+}) are replacing total Ca measurements using ISE technology.

Copper, magnesium, zinc and iron are trace elements in blood. They are quantitated readily by flame photometric, colorimetric or atomic-absorption techniques.

Organ Function Tests—The analyses of various blood or urine constituents, determination of metabolic excretion rates of exogenous compounds or endogenous metabolites and effect of exogenous stimuli on these parameters are used for evaluation of *in situ* activity and function of various organs. Organ function studies are performed in diseases associated with the liver, kidney, parathyroid, thyroid and pituitary gland, gastrointestinal tract, pancreas, adrenals and gonads. The principles and significance of the analysis used in such evaluations have been described also in other sections of this chapter.

Tests for *hepatic function* are based on bilirubin metabolism and excretion, carbohydrate metabolism (galactose tolerance test), plasma-protein changes (cephalin flocculation test and A/G ratio), abnormal fat metabolism, detoxification mechanisms (hippuric acid synthesis), excretion of injected substances [BSP], prothrombin formation and previously discussed enzyme levels.

Diseases of the liver are due to cellular alterations (hepatocellular) or obstructions to the flow of bile (obstructive jaundice). Hepatocellular liver disease can be chronic (postnecrotic cirrhosis, carcinoma) or acute (viral hepatitis, alcoholism, toxin- and chemical-induced).

The *cephalin flocculation test* is based on the flocculation of cephalin-emulsified cholesterol by γ-globulin. In normal serum an albumin-like protein will inhibit this reaction; in hepatic diseases, which produce abnormal γ-globulin or reduced albumin levels, the flocculation will occur.

The *detoxification mechanisms of the liver* can be evalu-

ated by intravenous administration of sodium benzoate and estimation of the benzoic acid metabolite, hippuric acid, in the urine. In hepatoparenchymal disease, a reduced capacity of the liver to form hippuric acid by conjugation of glycine and benzoic acid is observed.

The ability of the liver to excrete an injected dye is determined in the *BSP test;* the serum is analyzed for dye concentration at a suitable time interval after IV administration of 2 to 5 mg BSP/kg. Radioiodinated (^{131}I) Rose Bengal Sodium dye also has been used in dye-excretion studies with isotopic estimation of urine dye levels.

Kidney function tests are based on the determination of blood nonprotein nitrogen (urea, uric acid and creatinine), electrolytes, blood acid-base balance, routine urinalysis and the clearance of administered compounds in the urine. Most *clearance studies* are performed with substances that are not resorbed or secreted by the renal tubules: inulin, mannitol, sodium *p*-aminohippurate or ^{125}I-iothalamate sodium (sodium 5-acetamido-2,4,6-triiodo-*N*-methylisophthalamate). These are administered intravenously and the rate of urine clearance and glomerular filtration is estimated by analysis of the urine. The excretory capacity of the renal tubular epithelium can be determined by measuring the clearance rate of PSP. The dye is injected IV and the rate of its clearance in urine is determined. PSP is bound loosely to serum albumin and is removed rapidly from the blood by the renal tubules.

Sodium iodohippurate-(^{125}I), which is extracted almost completely from the blood on a single passage through the kidney, also has been used in renal function studies; a *renogram* or isotopic scan of both kidneys is performed. The test provides data on renal tubular secretion, renal vascular competence and renal evacuation and is primarily useful as a comparison of individual kidney function. It is important to note that 50% of kidney function can be compromised without any significant change in the routine renal function parameters.

Thyroid function tests usually measure the circulating levels of the thyroid hormones, and not the end-organ effect. The thyroid gland converts inorganic iodide to *thyroxine* (T_4) and *triiodothyronine* (T_3). T_3 and T_4 are stored in the colloid part of the gland as part of the thyroglobulin molecule. Hypothalamic *thyrotropin-releasing hormone* (TRH) mediates the release of the pituitary thyrotropin (*thyroid-stimulating hormone, TSH*). Excess levels of circulating T_4 depress, and low levels of T_4 increase, TSH release. TSH stimulates the proteolytic degradation of thyroglobulin to release T_4 and T_3, and increases organification of iodine. T_4 accounts for 90% of secreted thyroid hormones and exists in blood bound to *thyroxine-binding globulin* (*TBG*) or *thyroxine-binding prealbumin* (*TBPA*) or to albumin. T_3 is not protein-bound and has 5 to 10 times the biological potency of T_4 on a weight basis. Therefore, T_4 represents the major part of protein-bound iodine (PBI). The level of *free thyroxine* (FT_4), the active fraction in blood, is regulated by T_4 and T_3 release and the levels of binding proteins in blood and tissues.

The uptake of orally administered Na ^{131}I preparations by the thyroid gland can be estimated by isotopic scanning of the gland 24 hours after ^{131}I administration and is an index of glandular function (hyperactive, >50% uptake; hypoactive, <15%).

PBI determinations are based on the precipitation of protein-bound thyroxine, removal of inorganic iodine by basic- or anion-exchange chromatography, alkaline incineration to convert thyroxine to inorganic iodide and, finally, quantitation of iodide by reaction with arsenous acid and ceric ammonium sulfate. PBI is a good estimate of total circulating hormonal iodine. The normal range is 4 to 8 μg/100 mL serum.

T_4 can be determined by column chromatography in which it is separated and isolated by ion-exchange chromatography, and then analyzed colorimetrically. Nonisotope thyroid assays have been developed using fluorescence polarization methods for T_4 and free-thyroxin index. In the competitive protein-binding assay for T_4, serum T_4 competes with ^{125}I-T_4 for binding sites on a known amount of TBG. The ratio of bound to free ^{125}I is determined by adsorption of ^{125}I-T_4 not bound to TBG on an anion-exchange resin embedded in a polyurethane sponge or a porous dextran gel, and is a direct index of T_4 levels. The presence of mercurials, inorganic iodide or iodinated radiographic compounds in serum interferes with the T_4 column and PBI procedures. The competitive-binding procedure is affected by the presence of highly protein-bound drugs or changes in TBG levels in serum. The normal range of serum T_4 is 2.9 to 6.4 μg/100 mL by column and 3.0 to 7.0 μg/100 mL by binding assay. T_4 and PBI are increased in hyperthyroidism and the early stages of hepatitis. T_4 and PBI are decreased in hypothyroidism and nephrosis.

FT_4 also is determined in a competitive protein-binding assay in which ^{125}IT_4 and serum are incubated, and then dialysed to determine the percent dialyzable ^{125}I-T_4. FT_4 analysis is used in suspected abnormalities in protein-binding globulins. T_4 binding capacity of serum TBG, albumin and prealbumin can be determined after electrophoretic separation of these proteins.

T_3 analysis is determined by the resin-uptake test. The uptake of ^{125}I-T_3 by a resin is determined in the presence of the test serum. In hyperthyroidism, the primary TBG-binding sites are saturated and ^{125}I-T_3 is taken up by the resin. The resin uptake is decreased in hypothyroidism, and most of ^{125}I-T_3 is bound to TBG in serum. A *free thyroxine index* can be obtained by multiplying T_3 (resin) × T_4 (competitive binding) × 0.01. This product deviates from normal in the same direction as T_3 and T_4 in hyper- and hypothyroidism. This product is stable during euthyroidism in spite of changes in binding proteins; eg, a euthyroid patient on phenytoin therapy will show a decreased TBG and T_4 and increased T_3, but (T_4 × T_3) is normal. The indication of hyper- or hypothyroidism in the presence of abnormal amounts of TBG is observed in the (T_4 × T_3) product.

The determination of *TSH* by RIA appears to be the most useful test in discriminating patients with primary hyperthyroidism from the euthyroidism or hypothyroidism secondary to pituitary disease. Serum TSH is increased in the primary disease state.

The *PBI conversion ratio* is an estimate of the rate of conversion of inorganic iodide to PBI. Radioiodide-(^{131}I) is administered to the subject; after 24 hr, a sample of blood is obtained and the ^{131}I to P^{131}I is estimated by radiochromatographic procedures with ion-exchange resins (normal conversion, 13 to 42%).

Adrenocortical function is evaluated by estimation of serum or urinary 17-ketosteroids (17-KS) and 17-hydroxycorticosteroids (17-OH-CS) (androgen and corticosteroid metabolism), serum electrolytes (aldosterone metabolism) and blood adrenocorticotrophic hormone (ACTH) levels in the basal state, after stimulation with IM or IV ACTH, or after adrenal inhibition with dexamethasone. In the normal individual, ACTH will increase plasma cortisol and urine 17-OH-CS, and dexamethasone will suppress plasma cortisol. Metapirone, an inhibitor of 11β-hydroxylase, will cause selective secretion of compound S (11-deoxycortisol) by the adrenals in place of cortisol. Compound S will not inhibit the adrenal-pituitary feedback mechanism, the pituitary will secrete more ACTH and the adrenal will secrete more compound S. The determination of urinary 17-OH-CS or tetrahydro-compound S (THS) following metapirone administration is a good index of the functional integrity of the pituitary-adrenal axis; patients with virilizing adrenal hy-

Table IV—Reference Values[a]

Electrolytes		
Calcium	9.0–10.6 mg/dL	2.25–2.65 mmol/L
Chloride		98–109 mmol/L
CO_2 content		23–30 mmol/L
Magnesium	1.2–2.4 mEq/L	0.6–1.2 mmol/L
Phosphorus	2.5–5.0 mg/dL	0.81–1.62 mmol/L
Potassium		3.7–5.3 mmol/L
Sodium		138–146 mmol/L
Metabolites		
Bilirubin	0.1–1.2 mg/dL	1.7–20.5 μmol/L
Cholesterol	150–250 mg/dL	3.9–6.5 mmol/L
Creatinine	0.7–1.5 mg/dL (adults)	62–123 μmol/L
Glucose	60–95 mg/dL	3.33–5.28 mmol/L
Iron	50–165 μg/dL	9.0–29.5 μmol/L
Triglycerides	20–180 mg/dL	0.22–1.98 mmol/L
Urea nitrogen (BUN)	8–26 mg/dL	2.9–9.3 mmol/L
Uric acid	2.5–7.0 mg/dL	0.15–0.41 mmol/L
Proteins and enzymes		
Alanine aminotransferase	(ALT, SGPT)	5–40 U/L at 37°
Albumin	3.5–5.0 g/dL	35–50 g/L
Alkaline phosphatase	35–120 U/L at 37° (adults)	50–400 U/L at 37° (children)
Amylase	60–180 Somogyi Units	110–330 U/L
Aspartate aminotransferase	(AST, SGOT)	8–40 U/L at 37°
Carcinoembryonic antigen (CEA)	<2.5 ng/mL	<2.5 μg/L
Creatine kinase (CK)		10–180 U/L at 37°
Glutamyl transferase (GGT)		5–40 U/L at 37°
Lactate dehydrogenase (LDH)	60–220 U/L at 37°	(lactate → pyruvate)
Total protein	6.0–8.0 g/dL	60–80 g/L
Hormones		
Cortisol in plasma	7–20 μg/dL (at 8:00 AM) (200–550 mmol/L)	3–13 μg/dL (at 4:00 PM) (80–360 mmol/L)
Cortisol (free) in urine	20–90 μg/24 hr	55–248 nmol/24 hr
Follicle-stimulating hormone (FSH)	Adult males	Adult Females
	2–15 mIU/mL	Follicular phase 3–15 mIU/mL Ovulatory spike 10–50 mIU/mL Luteal Phase 3–15 mIU/mL Postmenopause 30–200 mIU/mL
17-Hydroxycorticosteroids in urine	3–10 mg/24 hr	
17-Ketosteroids in urine	5–15 mg/24 hr	(adult females)
	8–20 mg/24 hr	(adult males)
	0.1–3.0 mg/24 hr	(prepubertal children)
Luteinizing hormone (LH)	Adult males	Adult females
	5–25 mIU/mL	Follicular phase 5–30 mIU/mL Ovulatory spike 50–150 mIU/mL Luteal phase 5–40 mIU/mL Postmenopause 30–200 mIU/mL
Metanephrine in urine	<1.3 mg/24 hr	
Prolactin	1–20 ng/mL (males) (1–20 μg/L)	1–25 ng/mL (females) (1–25 μg/L)
Thyroxine (Ty)	5.5–12.5 g/dL (adults) (72–163 nmol/L)	7.8–16.0 μg/dL (newborns) (101–208 nmol/L)
Vanillylmandelic acid (VMA) in urine	<6.8 mg/24 hr	

[a] Serum specimens unless otherwise indicated.[18]

perplasia excrete excessive THS due to a 11β-hydroxylase defect.

Common, chemistry, reference values are listed in Table IV.[18]

Automated Analysis—The automation of analytical techniques used in blood and urine chemistry, hematology, blood typing and immunology has increased the productivity and accuracy of the clinical laboratory.[19] *Computerization* of the automated analytical system also has increased the rapidity of reporting test results, reduced clerical error and provided a unified and updated report of the laboratory tests for each patient.

In the SMA-12 (or SMA-20) Autoanalyzer (*Technicon*), a continuously operating, multiple-channel proportioning pump moves the samples, diluents and reagent streams. Air bubbles segment the flowing streams of samples and reagents, which then may flow through dialyzers to remove interfering substances, move them into chambers preset at desired temperatures and, finally, into detection devices (colorimeters, fluorometers, flame photometers, spectrophotometers). A serum standard is run simultaneously with the samples. The results can be read directly from a recorder or can be coupled into a digital computer output. Sequential, multiple analyses in the SMA-12 are accomplished by distributing the sample to 12 different analytical streams, so that all 12 analyses are in progress at the same time. The *SMA-12 profile* usually determines calcium, inorganic phosphorus, glucose, BUN, uric acid, cholesterol, total protein, albumin, total bilirubin, alkaline phosphatase, LDH and AST. The Mark X (*Hycel*), Ektachom 400 (*Kodak*), ACA (*Dupont*) and DSA-560 (*Beckman*) also are used in automated clinical-laboratory techniques.

Technicon recently developed the "capsule chemistry" analysis on the Chem 1 analyzer. Microaliquots of the sample (1 μL) and reagents (14 μL) are encapsulated within an inert fluorocarbon liquid. The resulting "test capsule" is introduced into a single, analytical flow path (composed of a solid fluorocarbon liquid, Teflon) where the sample is incubated, mixed, reacted and measured as a moving series of individual tests. The reactions are monitored at in-line detector stations for colorimetric and nephelometric measurements. On each sample 35 chemistries can be run sequentially.

The rapid growth of more-sophisticated chemistry analyzers increases the capacity of any clinical laboratory and is associated with small-specimen requirements incorporating batch analysis, profiles and stat capabilities. In addition to routine chemistry testing, the systems test for enzymes, immunoassay, therapeutic-drug tests, coagulation (fibrinogen, antithrombin III, plasminogen) and electrolytes. Techniques eliminating liquid requirements of other reagent systems are available from *Kodak* and *Ames* using dry reagents,

which are impregnated in pads on a strip or slide and read by a reflectance photometer.

Automated hematology and simultaneous determination of RBC, WBC, hemoglobin and hematocrit, MCV, MCH and MCHC can be performed on the SMA-7A (*Technicon*) Analyzer. The automated Technicon Hemalog system will provide data of SMA-7A and CCV (conductivity cell volume), prothrombin time, partial thromboplastin time and platelet count. *Automated leukocyte differential* was discussed previously under *Hematology*.

Urine

The formation of urine and its excretion are critical physiological activities of the body which provide a mechanism for the maintenance of a constant internal environment for all cells, tissues and organs. This internal ecology of the body is well-recognized and known as homeostasis. Inasmuch as the urine reflects what is occurring within the body, it offers a fluid which is an important source of information that is most useful as an aid in the definition of the states of health and disease. More specifically, the kidney, by means of urine formation

1. Regulates the body water.
2. Excretes metabolic waste products, many of which are of a nitrogenous nature.
3. Excretes toxic substances of both endogenous and exogenous origin.
4. Regulates the electrolyte equilibrium of the body by either excreting or retaining each specific ion.
5. Maintains the delicate balance of pH within the body by excretion of excess acid or excess base.
6. Provides an important route for the elimination of pharmaceutical agents and their breakdown products from the body.

Normal urine contains several thousand compounds most of which occur in minute quantities. Table V identifies some of the constituents of normal urine which are of particular significance.

Urine is studied quite widely as a means of identifying abnormalities associated with disease. The importance of such study is emphasized by the fact that the number of tests carried out on urine far exceeds those made on all other body fluids combined. Urine not only is important in providing information relating to kidney disease, but it may provide information relative to many other body activities. Information from urine studies is of diagnostic value in functional diseases of the kidney, liver, pancreas, blood, bone, muscle and the urinary, gastrointestinal and cardiovascular systems. Urine studies provide vital clinical information on electrolyte and water balance, acid-base equilibrium, intermediary metabolism, inborn errors of metabolism, drug abuse, intoxication, pregnancy and hormone balance. Most of these parameters have been discussed earlier and this section will be devoted to routine urinalysis.

Table V—Normal Constituents of Urine

Constituent	g/day	Constituent	g/day
Water	1400	Amino acids	2.1
Total solids	60	Purine bases	0.01
Urea	30	Phenols	0.03
Uric acid	0.4	Proteins (total)	0.025
Hippuric acid	0.9	Chloride (as NaCl)	12
Creatinine	1.2	Sodium	5
Indican	0.01	Potassium	2
Citric acid	0.8	Calcium	0.2
Lactic acid	0.2	Magnesium	0.15
Oxalic acid	0.03	Sulfur (total)	1.0
Nicotinic acid	0.00025	Phosphate (as P)	1.1
Allantoin	0.04	Ammonia	0.7

It is important to recognize that urine test information, like all other laboratory data, helps provide a picture of the whole body, but any single result requires interpretation to be most meaningful. It also should be recognized that negative results can be essentially as useful as positive results in a great many instances. The ready availability of urine is an advantage that makes it practical as a material for monitoring the course of the treatment of disease as well as for its recognition and definition.

Most urine examinations include observations with regard to the majority of the following—color, odor, turbidity, pH, protein, glucose (or reducing substances), ketone bodies (acetone), occult blood, bilirubin, urobilinogen, bacteria (culture or chemical tests), specific gravity and microscopic examination of sediment, including erythrocytes, leukocytes, casts, epithelial cells, crystals, bacteria, parasites and exfoliative cytology. A *"routine" urinalysis* varies in different institutions but ordinarily involves the inclusion of the majority of the above tests.

Urine for laboratory study should be collected in clean containers—preferably into a disposable unit (polystyrene tube) with a capacity of 15 mL which can be used for collecting, transporting, centrifuging and testing. Refrigeration is desirable for any specimen which is not tested within 1 to 2 hours.

If urine is to be transported through the mails or is to be held for a significant time at room temperature, it is desirable to add a urine preservative (formalin, methenamine, thymol, toluene) which will interfere with microbial growth in the specimen. Several proprietary urine preservative tablets are available. If urine is allowed to stand at room temperature, bacteria will grow in the specimen and cause degradation of many constituents. Frequently, the bacteria decompose urea into ammonium carbonate with a resulting increase in the alkalinity of the specimen. Formed elements, particularly casts and red blood cells, disintegrate in alkaline solution.

The majority of urine tests are done on random specimens but, in certain instances, it is necessary to have a 24-hr specimen for certain specialized analyses. For urine-sugar testing in diabetes detection, it is desirable to use a postprandial urine specimen (ie, after a meal). For protein tests, as well as chemical or culture tests for bacteriuria, the first morning specimen is preferred. Most laboratories use commercially available, standardized, reagent-impregnated strips ("dipstrips") or tablets (*Ames*) for routine urinalysis.

Instrumentation in Urinalysis—Automated urine-testing systems, semiautomated reagent-strip readers and a system which performs the complete urinalysis procedure have been developed. The strip reader is a reflectance photometer which measures urine pH, protein, glucose, ketones, blood, bilirubin, nitrate and urobilinogen. The IRIS AIM (*International Remote Imaging Systems*) measures urine specific gravity by refractometry, urine sediment by staining and classifies analytes, controlled fluid dynamics, video microscopy with an image processor, a chemistry system to read a standard dipstick by reflectance photometry, and color and appearance. These systems achieve standard results for routine urinalysis and increase accuracy and precision.

Volume—The normal volume of urine excreted during a 24-hr period is usually in the range of 1000 to 1500 mL. It is possible for a healthy person to modify the volume either by severe fluid restriction or by ingestion of excessive quantities of fluid. In certain disorders there is a change in urine volume. Urine-volume increases are identified as polyuria and are encountered in diabetes mellitus, diabetes insipidus and in certain stages of chronic renal disease. Urine volume is increased during diuretic therapy and with the ingestion or injection of large volumes of fluid. A decrease in urine

volume usually occurs in dehydration, water restriction and in acute or terminal renal disease. Extensive water loss from severe diarrhea or vomiting causes oliguria or decreased urine volume. Acute renal failure precipitated by shock, poisons or transfusion reaction may result in a complete absence of urine excretion or anuria. In the majority of instances urine study does not require volume measurements, but these are quite critical in severely ill persons where oliguria or anuria is present.

Specific Gravity-Osmolality—The urine density or specific gravity is related to the amount of solids excreted in a given volume of urine. In the majority of instances in healthy persons the specific gravity varies between 1.010 and 1.030 and is related to dietary habits of fluid and food ingestion and, secondarily, to the loss of fluid by other routes such as extensive sweating. The measurement of urine density or specific gravity is a part of "routine urinalysis," and as such provides information with regard to water and solids turnover in the body. The specific-gravity information alone is not nearly so important as it may be in conjunction with other observations. Thus, if dehydration is suspected, a specific gravity in the midrange of 1.015 would cast a doubt about dehydration unless there was a concurrent renal dysfunction.

The kidney possesses a remarkable ability to either form a concentrated urine or a very dilute urine ranging from a specific gravity of 1.001 to 1.032. This concentrating or diluting capacity is diminished in cases where there is a loss of renal function. In fact, one of the sensitive tests for measuring renal function involves the so-called dilution-concentration tests where fluid is administered or withheld, and the specific gravity of the urine is measured. With a serious loss of renal function, the kidney cannot excrete a urine in excess of 1.020 even with marked fluid restriction. In advanced renal disease the specific gravity of the urine may become "fixed" or constant in the range of 1.010 to 1.012 with all urine being of this specific gravity regardless of whether there is overhydration or dehydration.

Specific gravity is measured readily with a special hydrometer, called a urinometer. There is a correlation between the density of urine and its refractive index, and a special refractometer has been designed which gives readings in specific-gravity units on a single drop of urine.

Certain abnormal constituents of urine, such as glucose or protein, when present in high concentrations, will cause significant increases in specific gravity. Certain X-ray contrast media, when excreted in the urine, also will cause marked increases in specific gravity.

Urine specific gravity is only an indirect index of solute concentration, ie, 1 mole of urea will produce a lower specific gravity than 1 mole of glucose. Osmolality is a direct measure of the molal concentration of solutes in solution regardless of their molecular weight, ie, 1 mole of NaCl dissociates into 1 mole of chloride ion and 1 mole of sodium ion. Osmolality is determined in a direct-reading osmometer by comparing the freezing point of urine with that of a standard sodium chloride solution.

The kidneys normally excrete 800 to 1400 mOsm/kg (an osmol is that weight of any substance when dissolved in water depresses the freezing point 1.86°) of solutes per day. Man concentrates urine and eliminates the daily solute load at a maximum volume of 1200 mOsm/kg water. Urine osmolality is an inverse function of urine volume in the normal catabolic state. Urine volume is regulated by the antidiurotic hormone (ADH) and sodium excretion by the hormone aldosterone. Increased osmolality of body fluids stimulates, and increased dilution inhibits, the release of ADH. The major determinant of body-fluid osmolality is sodium. Sodium conservation is mediated through the renin-angiotensin-aldosterone axis. Determinations of plasma and urine sodium, and osmolality and urinary volume, are of diagnostic value in Addison's disease, vasomotor nephropathy (acute tubular necrosis), inapparent volume depletion, incomplete urinary tract obstruction and hepatorenal disease.

pH—Freshly voided urine usually has a slightly acid pH. The normal range is 5 to 8 and, essentially, this is also the abnormal pH range. The kidneys, by reason of excreting a urine of variable pH, provide a regulatory mechanism for the body to get rid of excess acid or alkaline waste products. Since the normal pH range and the abnormal pH range are comparable, the measurement of pH alone provides minimal information, but when used in conjunction with other information, it is a very useful urinary parameter. In conditions of acidosis, the urine is quite acid; in conditions of alkalosis, the urine pH is above 7. When metabolic or respiratory acidosis is suspected, an alkaline-urine pH result almost eliminates the possibility of acidosis. Conversely, if respiratory or metabolic alkalosis is suspected, the excretion of an acid urine indicates that alkalosis is likely not present.

Dip-and-read tests are used widely for pH testing, but pH-meter measurements are used less commonly. In certain situations involving kidney stone susceptibility, it is quite important to maintain a narrow range of urinary pH. For example, in cystinuria an alkaline pH is maintained to keep the cystine solubilized and to avoid as much as possible the crystallization of cystine into renal calculi. The maintenance of urinary pH is also important for optimum results in certain types of drug therapy.

Color—Urine normally has a yellow color, mostly due to urochrome; the color varies from pale straw to dark amber. Darker specimens usually have a high specific gravity. Occasionally, either normal or abnormal urine may show a color different from yellow. Bilirubin may cause fresh urine to be dark in color. In addition, urine which is allowed to stand darkens because of the oxidation of urobilinogen to urobilin. Red, reddish-brown or "smoky" urine usually is due to the presence of hemoglobin (hemoglobinuria), myoglobin (myoglobinuria) or red blood cells (hematuria). Porphyria is an uncommon cause of red coloration. Black urine can be caused by melanin, which may occur in the urine of patients with far-advanced malignant melanoma. An inborn error of metabolism, alkaptonuria, is characterized by the urinary excretion of homogentisic acid, which causes the urine to turn dark brown or black on standing. Many of the unusual colors occasionally found in urine are derived from exogenous sources, including both foods and drugs. Among these are the red color caused by beets, particularly in infants, the golden-yellow or orange-red color of metabolites of pyridium-like drugs or azo drugs and the green or blue color from methylene blue.

Odor—Normal, freshly voided urine has a faint aromatic and characteristic odor, which is more intense in concentrated specimens. If the urine is allowed to stand, the odor becomes strongly ammoniacal and unpleasant because of bacterial destruction of urea. Freshly voided urine having a foul odor indicates severe infection. A sweet, fruity odor may be due to ketones.

Appearance—Freshly voided urine is usually clear. On standing, a precipitate may form which usually consists of amorphous urates if the urine is acid or calcium and magnesium phosphates if the urine is alkaline. The formation of a precipitate is more likely to occur if the urine is refrigerated. Most specimens will become clear again if they are warmed gently to room temperature. Large quantities of mucus, cells, leukocytes or bacteria may cause cloudiness. Protein usually does not cause cloudiness.

Protein—A small amount of protein is present in the urine obtained from healthy subjects although the quantity is not sufficient to give a positive reaction with the tests

commonly used for the recognition of protein in urine. The majority of the 25 to 50 mg of protein that is excreted daily is microprotein (low-molecular-weight polypeptide), with properties quite different than those of albumin and globulin, which are the principal proteins of the blood serum. Albumin and globulins do occur in the normal urine in minute concentrations.

Plasma proteins, hemoglobin, abnormal Bence-Jones protein and proteins (nucleo-, phospho- and glyco-proteins) derived from leukocytes and mucus may be present in urine in nephritis, nephrosis, lesions of the urinary tract, GI dehydration and renal congestion. Abnormal amounts of protein in the urine may be recognized by either precipitation or colorimetric tests. The precipitation depends on the heat coagulation of the protein or on the chemical precipitation of the protein. The most popular of the heat-precipitation tests is the heat-and-acetic acid test in which a tube of urine is heated to boiling after the addition of a drop or two of acetic acid. Sulfosalicylic acid is employed commonly in chemical precipitation tests and, in this test, equal quantities of 3% sulfosalicylic acid and urine are mixed in a test tube and the mixture examined for turbidity indicative of precipitated protein.

Colorimetric tests for proteins involve *dip-and-read* type of systems and are based on the *protein error* of indicators. Certain indicators have a point of color change which is different in the presence of protein compared to the same system in the absence of protein. Thus, by buffering the indicator tetrabromophenol blue on this dip-strip at a specific pH, it is possible to have a yellow color in the absence of protein and a green or blue color in the presence of protein. This test, Albustix (*Ames*), not only indicates the presence or absence of protein in the urine but also can be made to indicate the approximate amount of protein. Strongly alkaline or fermented urines will give false-positive results. The sensitivity of the colorimetric method is such that quantities of 10 to 20 mg of albumin per 100 mL of urine are recognized with confidence.

A positive test for protein in the urine may have any one of several meanings, and it is only when this information is related to other observations that it has optimum value. Proteinuria may be benign and appear following strenuous exercise or simply as a result of standing (orthostatic proteinuria). Protein frequently occurs in the urine during pregnancy and in some instances this is benign, but in other cases it indicates renal complications. Transient proteinuria may occur following severe infections, high fever, exposure to cold and in congestive heart failure. Proteinuria may be an early and sensitive indicator of renal disease and may indicate an abnormality prior to other signs and symptoms of renal impairment in the glomerulus or tubules. In the majority of instances there is not a correlation between the amount of protein in the urine and the severity of the renal disease.

Patients with severe nephrosis may lose up to 25 g of protein per day. Such a marked loss of protein causes a decrease in plasma protein concentration with an accompanying edema. In both chronic and acute glomerulonephritis there is protein in the urine. Tumors of the kidney and renal infection usually will have an accompanying proteinuria. Bence-Jones protein is a unique protein which occurs in the urine of about 50% of patients with multiple myeloma. It has the unusual property of precipitating between 50 and 60° and dissolving at higher temperatures.

Glucose (Reducing Substances)—Glucose normally occurs in urine in such low concentration that it escapes detection by the usual testing methods. The urine of untreated or poorly controlled diabetic patients characteristically contains easily detectable amounts of glucose. A positive test for glucose in urine usually suggests hyperglycemia and the diagnosis of diabetes mellitus; further studies, such as the glucose tolerance test to confirm the diagnosis, are indicated. Glycosuria also may occur when the renal tubules fail to reabsorb glucose normally, and glucose appears in the urine despite normal blood glucose levels, in contrast to true diabetes.

Glucose is the sugar almost always found in urine; however, lactose, galactose, levulose, sucrose and pentoses may be encountered. These other sugars are identified by paper chromatography, selective fermentation, polarimetry, special chemical tests or the formation of their osazones. Other reducing substances occur in urine and may cause falsely positive reducing reactions for glucose. Examples are ascorbic acid, glucuronides, many drugs, homogentisic acid and the preservatives formalin and chloroform.

The traditional test for glucose in urine (Benedict's test) relies on the reduction of cupric ions in alkaline solution to reddish-orange insoluble cuprous oxide. The copper is reduced totally by large amounts of glucose and results in a brick-red sediment with no remaining blue color. Lesser concentrations form green- to rust-colored solutions with some red sediment. A modification of this test, Clinitest (*Ames*), is available in tablet form. The tablet contains copper sulfate, anhydrous sodium hydroxide, citric acid and sodium carbonate. When added to dilute urine, the tablet dissolves and generates enough heat and effervescence to yield results comparable with the Benedict test.

A specific but extremely simple enzyme test for glucose is available—Tes-Tape (*Lilly*), Clinistix (*Ames*) and Multistix (*Ames*). Reagent strips are impregnated with glucose oxidase, peroxidase and orthotolidine. When dipped into a solution of glucose, oxidation occurs and hydrogen peroxide is formed which oxidizes orthotolidine to a blue color. This test is more sensitive than Clinitest, but is not as reliable for estimating the concentration of glucose. The enzymatic test is specific and thus useful in determining whether or not a reducing substance is glucose. Diastix (*Ames*) is a specific urine glucose test using glucose oxidase, which also indicates the quantity of glucose present.

Ketone Bodies—The ketone bodies acetone, acetoacetic acid and beta-hydroxybutyric acid are present in the urine when fats are metabolized incompletely. Ketonuria is seen most commonly in poorly controlled diabetes and indicates ketonemia and diabetic acidosis. Other causes for ketonuria are starvation, fever, protracted vomiting and Von Gierke's disease. Ketonuria also occurs following anesthesia. Acetoacetic acid and acetone produce a distinctive purple color when treated with a mixture of sodium nitroprusside, ammonium sulfate and concentrated ammonium hydroxide. A similar reagent is available in tablet form (Acetest, *Ames*). A drop of urine is placed on the tablet; if ketones are present, a lavender to deep-purple color develops in 30 sec. The color intensity indicates the concentration of ketones. The reagent strip Ketostix (*Ames*), used as a dip-and-read test on urine or serum, contains the same reagents, which are available on Multistix (*Ames*) and other multiple reagents as well. These tests will detect 5 to 10 mg of acetoacetic acid per 100 mL of urine.

Phenylpyruvic Acid—Phenylketonuria (or PKU) is an inborn error of metabolism in which the normal conversion of phenylalanine to tyrosine in the body does not occur and there is a buildup of phenylalanine concentration in the blood. This metabolic disorder causes mental retardation. A portion of the phenylalanine is excreted by the kidneys into the urine and in the process is converted to phenylpyruvic acid (or phenylketone). If this genetic disorder is discovered soon after birth, it is possible to place the infant on a diet very low in phenylalanine-containing proteins and thus minimize the phenylalanine buildup in the body, averting

the serious mental retardation which ordinarily is seen in the untreated PKU patient.

Recognition of PKU can be made by the use of a test for phenylpyruvic acid using a dip-and-read reagent composition containing ferric ions. This test, Phenistix (*Ames*), can be used on urine from all newborn babies. A positive reaction gives a green color, whereas a normal infant's urine gives a pale-ivory or yellow color to the strip. PKU also can be recognized by employing a chemical or microbiological test for elevated phenylalanine in serum, as discussed under *Amino Acids*.

Bilirubin—Bilirubin is found in the urine of patients with hepatitis or obstructive jaundice but not in patients with hemolytic jaundice. Tests for bilirubin and urobilinogen combine to give excellent information in the differential diagnosis of jaundice. Tests for bilirubin are of two kinds; oxidation tests form a green color of biliverdin from bilirubin usually using ferric chloride as the oxidative reagent, and diazotization tests form colored compounds when bilirubin reacts with diazonium salts in a strongly acid medium. Most oxidation tests adsorb the bilirubin onto barium sulfate or similar material before the addition of Fouchet's reagent. The tablet test Ictotest (*Ames*) is the most sensitive diazo test and it uses an absorption mat to concentrate the bilirubin from 5 drops of urine. A reagent tablet is added to the moist spot on the mat and 2 drops of water are added to dissolve the effervescent reagent and wash some of it off the tablet onto the mat where the reaction takes place. A blue or purple color on the mat around the tablet in 30 sec indicates the presence of bilirubin. In addition, a dip-and-read test composition also based on the diazo reaction has been incorporated into the multiple urinalysis reagent strips, Bili-Labstix and Multistix (*Ames*). It is less sensitive than the tablet test, but its convenience allows it to be used in routine urinalysis quite readily. An incidence of approximately 0.1% positives on health-screening population groups, 0.2% on clinic patients and 0.9% on hospitalized patients has been reported.

Urobilinogen—Bilirubin in the bile is reduced to urobilinogen by bacteria in the lower intestine. A portion of the urobilinogen is reabsorbed from the intestine into the blood. A portion of this urobilinogen is excreted into the urine by the kidney and the balance is re-excreted via the bile into the intestine. Although the quantity of urobilinogen in the urine is quite small, it is an important indicator of liver function and red-blood-cell catabolism.

If there is an obstruction to bile flow such as in obstructive jaundice, the amount of urobilinogen formed and reabsorbed into the blood and excreted in the urine is decreased. With impairment of liver function, the excretion of urobilinogen in the bile is decreased, the blood concentration increases and there is a corresponding increase in urinary urobilinogen excretion. Actually, the increase in urinary urobilinogen is one of the most sensitive tests for impaired liver function and this test may indicate an abnormality when all other tests of liver function remain unchanged from normal.

In hemolytic diseases where there is an increased rate of hemoglobin breakdown, the amount of bilirubin formation is increased with a corresponding increase in urobilinogen formation and excretion in the urine. The concentration of urobilinogen in urine can be established by the use of a dip-and-read test which uses the interaction of urobilinogen and *p*-dimethylaminobenzaldehyde (Urobilistix, *Ames*).

Hematuria, Hemoglobinuria and Myoglobinuria—Hematuria refers to a condition in which intact red blood cells appear in the urine. This condition is indicative of a specific defect in the microscopic functional unit (the nephron) of the kidney or it may be indicative of bleeding in the kidney, the ureter, the bladder or the urethra. In the female there may be variable numbers of red blood cells in the urine during menstruation.

Hemoglobinuria is a condition in which free hemoglobin is present in the urine without red blood cells. This may be caused by intravascular hemolysis as a result of a transfusion reaction or by poisoning or toxins. The free hemoglobin in the plasma is excreted by the kidney into the urine. In some situations actual total hemolysis of the red cells occurs after they have entered the urine. This occurs particularly with alkaline urines.

Myoglobin is the red respiratory pigment of muscle. This pigment is quite comparable to hemoglobin in its composition and chemical reactions. Myoglobin may be liberated from muscle cells in certain types of injury and, in such cases, will circulate in the plasma and be excreted in the urine. There are also certain genetic muscle disorders in which myoglobin is lost from the muscles and appears in the plasma and subsequently in the urine.

Chemical tests for red cells, free hemoglobin and myoglobin are based on the peroxidase-like activity of hemoglobin or myoglobin. When a chromogen mixture such as orthotolidine and peroxide is exposed to this peroxidase activity, it will interact rapidly to generate an intense blue color. A dip-and-read solid state system is available which is called Hemastix (*Ames*). This specific composition uses cumene hydroperoxide as the peroxide. The same dip-and-read test for occult blood is incorporated as a component part of multiple, urine dip-and-read tests, eg, Multistix (*Ames*).

Microscopic Examination—Ordinarily, urine contains a number of formed elements or solid structures of microscopic dimensions. These are studied readily by centrifuging 10 to 15 mL of urine, pouring off the supernatant and resuspending the sediment in the drop or so of urine which remains in the tube. This suspension of sediment is placed on a microscope slide and viewed with low-power magnification. Specific structures can be studied with higher magnification. The urinary sediments can be classified into unorganized (chemical substances) and organized (cells and casts) constituents.

In an alkaline urine, amorphous or crystalline ammonium-magnesium phosphates, calcium carbonate or oxalate crystals and ammonium urate may occur normally. Amorphous or crystalline urates, uric acid and calcium oxalates normally are seen in acid urines. The presence of tyrosine, leucine or cystine crystals is associated with various diseases. Chemical crystals are identified by solubility in acid and/or alkali, colorimetric reactions and crystalline structure.

The urine sediment ordinarily contains residues of epithelial cells, crystals and an occasional red or white blood cell. Increased numbers of erythrocytes are seen where there is bleeding into the urinary tract. If the red cells are formed into a red-cell cast, it is suggestive that bleeding has occurred at the glomerular level. An increased number of leukocytes is suggestive of infection and inflammation of the kidney. Casts are microscopic concretions which have the form of a tubule; they have a matrix of precipitated protein and, depending on their appearance, may be identified as hyaline, granular, waxy or red-cell casts. Renal-failure casts are larger and are associated with severe necrosis of the kidney.

Numerous crystals, mucus fibers, bacteria, yeast cells, spermatozoa and parasites (*Trichomonas vaginalis*) may be indentified in the urine sediment. The majority of these crystals do not have any unusual significance but in certain disorders may be indicative of crystal deposits in kidney tissue or predisposition to formation of calculi.

Tissue cells can be recognized in urine sediment. This provides an excellent means of detection and diagnosis of cancer of the lower urinary tract when the sediment is fixed in alcohol and stained by the Papanicolaou procedure. Ex-

foliative cytology of urine may be applied as a routine to all urology patients. In one large clinic the number of positive cases found among urology patients was almost 5%, which is a much higher return of positive results than is obtained with routine staining of cervical smears.

Bacteria—Freshly voided specimens of urine ordinarily contain a few microorganisms, which primarily represent bacteria picked up from the external genitalia. There are fewer contaminating organisms in a *clean-catch* specimen, which involves extensive washing of the external genitalia prior to collection of the specimen. A specimen collected at the midpoint of urination or a "midstream" specimen ordinarily has more organisms than a clean-catch specimen, but fewer than a so-called random specimen. When there is an infection of the kidney or urinary tract, the number of organisms in the urine is increased markedly. Ordinarily, if the urine contains 100,000 or more organisms per mL, the result strongly suggests the presence of an active infection. Infection of the urinary tract with accompanying bacteriuria is relatively common in young girls and women. Quite often the condition is asymptomatic and is recognized only as a result of a study of the urine. If bacteriuria is not treated, it may lead to serious renal injury.

If there is a very large number of bacteria in the urine, the specimen actually may be turbid. This can be recognized by gross visual inspection of the urine. Bacteriuria also can be recognized by microscopic examination of the urine sediment particularly if there is a large number of organisms present. The most widely employed procedure for recognizing bacteria involves plating a specimen of diluted urine on a culture plate, incubating it and counting the number of colonies. A more convenient approach to this same measurement involves the use of a microscope slide which is coated with nutrient agar. Such a slide, when dipped in a urine specimen and then incubated, will indicate the presence or absence of bacteriuria and also the approximate count.

Methods to determine the presence of significant numbers of bacteria in urine samples are available on various automated systems.[20] The Bac-T-Screen (*Marion*) system is a dispensing and filtering system used with a straining process to detect the presence of bacteria on special filter cards by noting the color change on the card. Analysis on the Abbott MS-2 performed by photometric monitoring of bacterial growth changes the light transmitted in a broth culture over a period of time. A decrease in the light transmission due to turbidity or color identifies a positive specimen. The Lumac Biocounter M2010 measures bacterial adenosine triphosphate (ATP) in urine by the bioluminescence produced in a luciferin-luciferase system. Once these rapid techniques are performed to determine which specimens have increased bacteria, further identification and sensitivity testing are performed. Chemical tests for the metabolic activity of bacteria have been used in studying bacteriuria. The most popular chemical test is that for nitrite. Ordinarily, all urine specimens contain nitrate, but do not contain nitrite. If *E coli*, or certain other organisms, are present in sufficient numbers, they will reduce the nitrate to nitrite.

Calculi—Knowledge of the composition of renal and bladder calculi ("stones") is essential in planning the therapeutic regimen for such diseases. Mixed calcium phosphate and oxalate stones usually occur over the entire urine pH range. Uric acid, cystine and calcium hydrogen phosphate calculi generally are associated with acid urines, while magnesium ammonium phosphate calculi usually occur in alkaline urine. Hyperexcretion of one of the calculi components, pH, renal blockage and the presence of foreign objects in the urinary tract are the most probable causal factors in the formation of renal calculi. Calcium oxalate stones are the most common type. The chemical content of the stones is established by routine qualitative analysis for calcium, magnesium, ammonium, phosphate, carbonate, oxalate, uric acid and cystine. Subsequent confirmation by optical crystallography, X-ray diffraction and infrared spectroscopy is also used in the characterization of the physical properties of the calculi.

Feces

Normal feces consists of undigested food remnants, products of digestion, bacteria and secretions of the gastrointestinal tract. *Macroscopic, chemical* and *microscopic* determinations are performed routinely. The normal quantity of feces is about 200 g/day. The brown color is due to the reduction of bilirubin to urobilinogen and then to uribilin (stercobilin); bilirubin is not normally present in feces, but porphyrins and biliverdin (a component of meconium) are excreted during the first days of life. Bilirubin can be detected by tests previously described for bile pigments.

Color changes in the stool can be the result of dietary intake or diagnostic for biliary obstruction and gastrointestinal bleeding.[21] Patients with steatorrhea and malabsorption may show a yellow bulky stool containing fat and gas. The feces is clay colored when bile is prevented from entering the gut. A red or black stool can occur when excessive doses of anticoagulants, phenylbutazone or salicylates are taken, producing bleeding in the gastrointestinal tract. Substances which interfere with the coloration of the stool include antacids (whitish or speckling), bismuth salts (black), iron salts (black), pyridium (orange), senna (yellow to brown) and tetracyclines (red).

Fecal urobilinogen can be determined colorimetrically by reduction of urobilin to urobilinogen with alkaline ferrous sulfate, and then reaction with acidified *p*-dimethylaminobenzaldehyde (Ehrlich's reagent). It is increased from a normal range of 40 to 280 mg a day to 400 to 1400 mg in hemolytic jaundice (dark brown stool), and is decreased in obstructive jaundice (clay-colored stool).

Porphyrins and porphyrinogens do not arise from hemoglobin catabolism, such as bilirubin, but are by-products of the synthesis of heme. Increases in fecal and urinary elimination of coproporphyrin, uroporphyrin and protoporphyrin are valuable diagnostic aids in distinguishing the various hepatic and erythropoietic porphyrias. Fecal coproporphyrins (CP) and coproporphyrinogens (CPP) are determined after extraction, conversion of CPP to CP by iodine and triple-point spectrophotometric estimation at 380, 401 and 430 nm to correct for interfering substances (also see *Urinalysis*).

Fecal occult blood is detected readily by the *o*-tolidine, benzidine, guaiac or diphenylamine tests; this is valid only if the patient has been on a meat-free diet for 3 days. Guaiac and diphenylamine are preferred due to the carcinogenic potential of the other two chemicals.

The Hemoccult test kit (*SmithKline Diagnostics*) uses an impregnated guaiac paper slide for detecting occult blood, which is a useful screening test for colon cancer. Two slides are prepared each day for 3 days from different parts of the same stool while the patient is on a meat-free high-bulk diet. Interfering substances include aspirin, indomethacin and corticosteroids, because they can produce bleeding, and Vitamin C, which interferes with the oxidation reaction of the test. If bleeding occurs high in the GI tract, the blood is digested and converted to acid hematin; 50 mL of blood in the feces will cause melena (black stool). Bleeding from the lower GI tract is apparent from red streaking of stools. The use of ^{51}Cr-tagged erythrocytes has been used to quantitate and locate the source of gastrointestinal bleeding. The subject's red cells are mixed with an isotonic ^{51}Cr solution and then reinjected intravenously. If bleeding occurs, the ^{51}Cr-

isotope content of the feces will be increased. Location of the hemorrhagic area also can be approximated by an isotopic scan of the abdominal area.

The presence of excessive quantities of *mucus* is usually indicative of dysentery, colitis or other inflammatory processes in the intestinal mucosa. Strongly alkaline or acidic reaction in the feces is indicative of excessive quantities of protein or carbohydrate in the diet, respectively.

Quantitative determination of *fecal nitrogen* is useful in analysis of pancreatic function. In pancreatic disease, increases in fecal nitrogen will occur as a result of decreased secretion of pancreatic proteolytic enzymes. The normal individual will excrete 4 to 13% of ingested nitrogen in the feces; in chronic pancreatitis, 9 to 30%. Fecal nitrogen can be determined by the Kjeldahl digestion procedure.

Fecal fat is present in the form of triglycerides of fatty acids (neutral fat), free fatty acids (FFA) and soaps. Fat determinations are based on the solubility of neutral fat and FFA in ether; the soaps are insoluble in ether and have to be acid-hydrolyzed to their respective FFA prior to extraction. Neutral fat will liberate FFA only on alkaline hydrolysis. The FFA, isolated from the above fractionations, are then determined by titrimetric, colorimetric or gas-chromatographic procedures.

Determinations of blood, urine and fecal ^{125}I after oral administration of an iodinated glyceryl trioleate or ^{125}I-oleic acid preparation is an index of *pancreatic, biliary and intestinal absorptive function* and correlates with *fecal fat excretion*. The bile must emulsify the ^{125}I-triglyceride prior to enzymatic hydrolysis by pancreatic lipase to yield FFA-^{125}I, which subsequently is absorbed and metabolized. An increased amount of ^{125}I in the feces is associated with pancreatic diseases (cystic fibrosis with achylia), obstructive jaundice, malabsorption disease (sprue, celiac disease) and steatorrhea. The latter entity can be differentiated as to a pancreatic lipase or intestinal absorptive defect. In the "absorptive" disease, increased excretion of ^{125}I is seen after administration of ^{125}I-triolein or oleic acid. In the pancreatic defect, adequate absorption of ^{125}I oleic acid occurs but fecal ^{125}I is increased after the triolein meal.

A *microscopic examination* of emulsified feces includes analysis for the presence of crystals, food residues, body cells, bacteria and parasites. Crystals of triple phosphate, calcium oxalate, fat and cholesterol, starch granules, vegetable fibers and neutral fat globules are normally present. Octahedral needle-shaped crystals (Charcot-Leyden crystals) are present in parasitic infestation and mucous colitis. Excessive quantities of fat or starch are seen in malabsorption disease.

Adult, larval or ova phases of parasites may be encountered in the feces. The most common parasitic infestations are caused by *cestodes* (tapeworms), *trematodes* (flukes), *nematodes* (roundworms) and *protozoa* (amoeba) (see *Microbiology*).

Toxicology

The determination of drug or chemical concentrations in biological fluids is an important aspect in diagnosing and treating the toxic syndrome induced by various agents in acute or chronic drug-abuse situations or in chemical poisoning.

Barbiturates, glutethimide, methaqualone, chlordiazepoxide, diazepam, diphenhydramine, ethchlorvynol, morphine, phenothiazines and salicylates are encountered in drug-abuse situations. Preliminary screening of serum or urine samples for drug substances is accomplished by TLC procedures. The analysis of serum or urine levels of intact drug or its metabolites usually is performed by extraction of the sample with an organic solvent, separation by gas-liquid (GLC), or high-performance liquid (HPLC) chromatography, and quantitation by spectrophotometric, fluorometric or electrochemical techniques. The interpretation of the serum-concentration data in relation to clinical significance and toxicology must not be limited to numbers.

In acute drug overdosage the time of drug ingestion, time of blood or urine sampling and severity of clinical symptoms or time of death must be interpreted in reference to data on the absorption, tissue distribution, metabolism and elimination of the drug and its metabolites. The specificity of the chemical assay as to interference from other drugs or metabolites of the parent drug must be considered. The combined techniques of GLC or HPLC and mass spectrometry confirms the identity of specific drugs in biological matrices. The extent of absorption of many drug substances is not related directly to the dose when large amounts of a drug are ingested, in comparison to the therapeutic dose.

The tissue-distribution and metabolic rates can be affected by large drug overdoses in which renal or hepatic failure is encountered. The plasma-elimination rate also can be affected, and it is important to recognize the change in elimination kinetics and to be aware of the nature of plasma elimination as defined by a mono-, bi- or polyexponential elimination curve. The drug overdose usually involves several drug substances and the chemical, metabolic and pharmacological aspects of drug interaction must be considered.

The methodology for the analysis of drugs in biological fluids or tissues can be found in the books listed in the *Bibliography*. Analysis for serum *barbiturate* levels will be described in this section as a specific example of the analytical methodology.

Serum is extracted at pH 6.5 with chloroform; the chloroform extract is washed with pH 7.0 phosphate buffer and extracted with 0.45*N* NaOH. The UV spectrum of the alkaline aqueous layer is determined at pH 13 and 10.5. The UV spectra are characteristic and distinguish barbiturates, *N*-methylbarbituric acids and thiobarbiturates. The barbiturates also can be detected by acidifying the alkaline layer, extracting with chloroform and spotting this organic extract on a silica-gel TLC plate. Sequential spraying of the plate with $KMnO_4$, $HgSO_4$ and diphenylcarbazone will show R_f values and color reactions typical of the various barbiturates. Blood barbiturates can be determined more accurately by a GLC procedure in which the retention times are used to identify the specific barbiturates. The degree of severity of clinical symptoms has been correlated with blood barbiturate levels. Comatose, areflexic signs are observed at 5.0 mg% amobarbital, 2.0 mg% pentobarbital, 8.0 mg% phenobarbital and 1.5 mg% secobarbital.

Opiates, amphetamines, barbiturates and methadone can be detected rapidly by *"homogenous" enzyme assay*.[22] In this procedure, the addition of drug antibodies to a conjugate of drug and lysozyme results in the inhibition of lysozyme activity. The addition of free drug to this reaction mixture increases the enzyme activity in proportion to the amount of free drug added. The sensitivity of this type of assay is 0.1 μg/mL of amphetamine and barbiturates, 0.5 μg/mL of methadone, 0.3 μg/mL of opiates and 1.0 μg/mL of benzoylecgonine, a cocaine metabolite. This assay is applicable to large drug-screening programs.

Electron-spin-labeling techniques also can be employed on large-scale drug-screening programs. In this procedure known amounts of drug antibodies are mixed with drug labeled with a stable nitroxide radical (spin-label) and with the specimen to be analyzed. Due to the competition for antibody between spin-labeled drug and drug in the specimen, the spin-labeled drug becomes detached from the antibody and can be detected by electron-spin resonance spectroscopy. This procedure is 1000 times more sensitive than TLC.

Blood-alcohol levels may be determined by aeration, distillation, gas chromatography or specific enzymatic analysis with alcohol dehydrogenase. In the chemical techniques the blood sample is either oxidized or distilled into a dichromate-sulfuric acid mixture; the excess dichromate is then determined by titration with potassium iodide or methyl orange-ferrous sulfate solutions or by colorimetric analysis. The gas-chromatographic and enzyme procedures are specific for ethanol, whereas the chemical techniques are influenced by other volatile or oxidizable substances in the blood. The enzymatic method is based on the reaction of ethanol and NAD in the presence of alcohol dehydrogenase to form acetaldehyde and NADH; the acetaldehyde is removed with semicarbazide and the NADH formed in the reaction is estimated spectrophotometrically at 340 nm. Ethanol levels of >0.10% are indicative of intoxication and apparent psychomotor disturbance. Levels of 0.40 to 0.50% are associated with medullary and diencephalic disturbances such as tremors, coma, respiratory depression, peripheral collapse and death.

Specific analysis of heavy metals is best performed by atomic-absorption spectroscopy. Analyses for arsenic, beryllium, bismuth, copper, iron, lead, lithium, mercury, nickel, thallium and zinc are encountered frequently in the toxicology laboratory. *Blood lead* is determined by forming a lead-dithiocarbamate chelate in the presence of ammonium pyrrolidinedithiocarbamate and extracting the chelate into methyl isobutyl ketone for subsequent atomic-absorption analysis. A lead concentration of >60 μg/mL in children usually reflects significant absorption and accumulation of lead and is interpreted as an indicator of lead toxicity (plumbism).

Increased lead exposure will result in a decrease in delta-*aminolevulinic acid* (*ALA*) conversion to porphobilinogen by ALA-dehydrase in heme synthesis. ALA blood levels will increase to the point that ALA is excreted in the urine. Determination of urinary ALA is performed by removing urine porphobilinogen and urea by ion-exchange chromatography, reacting ALA with *p*-dimethylaminobenzaldehyde and determining the chromogen colorimetrically. Urinary ALA levels >2.5 mg/100 mL are unacceptable in children and industrial lead workers. Urinary ALA levels are not as sensitive an indicator of lead toxicity as blood lead, but they can be used to monitor prophylactic treatment procedures.

Cholinesterase determinations are of value in the diagnosis of suspected cases of organophosphate or carbamate pesticide poisoning. Two types of cholinesterase are found in tissues. True cholinesterase is found in RBC and nerve tissue and exhibits a specificity for acetylcholine substrate. Pseudocholinesterase is found in plasma and has a greater affinity for hydrolyzing butyrylcholine and other esters. The organophosphate and carbamate insecticides inhibit both enzymes. The activity of the plasma enzyme is inhibited more rapidly than the RBC cholinesterase, and recovers more rapidly due to synthesis of new enzyme by the liver. The recovery of the erythrocyte enzyme is slow and is governed by red-cell turnover rate. Cholinesterase activity usually is determined by measuring changes in pH after the incubation of plasma or RBC lysates with acetylcholine. The normal range of this enzyme is 4.5 to 10.9 (plasma), 3.4 to 5.7 (whole blood) and 6 to 10.5 (RBC) units/mL.

Gastric Analysis

The chief constituents of gastric juice are hydrochloric acid, gastric proteases (pepsin and gastricsin), hematopoietic factor (intrinsic factor and vitamin B_{12} binders), gastric hormones and mucosubstances (aminopolysaccharides, mucopolyuronides, mucoids and mucoproteins). Tests for *gastric function*[23] usually are performed on gastric juice samples collected by direct intubation into the stomach. The fasting content (normal, <100 mL) of the stomach is removed and gastric secretion is collected in the basal state, or after stimulation by the oral administration of caffeine-benzoate or alcohol, or parenteral administration of histamine, insulin or the hormone pentagastrin. Samples are collected by continuous aspiration and analyzed for acidity and gastric protease activity at various time intervals. The extent of recovery of total juice can be estimated by oral, nonabsorbable indicators (polyethylene glycol-^{14}C, phenol red and ^{125}I-HSA) instilled into the stomach prior to the aspiration. The recovery and specific concentration of these indicators in gastric juice is an index of gastric secretory volume, completeness of collection and gastric emptying rate.

Gastric juice is a heterogeneous mixture of clear juice and flocculent, clear mucus. The *color* of the juice should be noted as to the appearance of blood, bile and excessive quantities of mucus. The *acidity* can be determined by a simple pH measurement and conversion to mEq H^+ or by titration of centrifuged gastric juice to pH 3.5, 4.5 and 7.4, the respective end-points for free acid (HCl), protease activity and physiological neutrality. The *basal acid output* is about 1 mEq/hr in normal subjects and 2 to 4 mEq/hr in duodenal ulcer patients. The *peak acid output* (*PAO*) after histamine stimulation is 10 to 20 mEq/hr in normals and 40 to 50 mEq/hr in duodenal ulcer; PAO following pentagastric stimulation is similar to histamine. Gastric acid secretion is decreased in atrophic gastritis, gastric carcinoma and certain types of gastric ulcer. Hypersecretion is seen in duodenal ulcer, Zollinger-Ellison (ZE) syndrome and hyperparathyroidism.

In situ measurements of pH may be made with a *Heidelberg capsule apparatus*. In this technique the subject swallows a small pH-sensitive capsule (transmitter); radiowaves are transmitted from the capsule to a sensing device (receiver), and the signals are recorded as a function of pH. The normal pH of the stomach is 1.2 to 1.8.

Tubeless gastric acidity analysis is performed by oral administration of Diagnex Blue (*Squibb*), a carbacrylic ion-exchange resin reacted with azure blue dye. The hydrogen ions in the gastric juice exchange with the dye on the resin; the dye is absorbed and then excreted in the urine. The dye concentration in the urine is a function of gastric acidity. The normal value is >0.6 mg of dye in the urine 2 hours after administration.

The principal gastric proteases are *pepsin and gastricsin;* pepsinogen is a precursor which is converted to active pepsin by free HCl and by an autocatalytic process. *Total gastric protease activity* is determined on hemoglobin or radioiodinated human serum albumin (RISA) substrates at pH 1.8 to 3.1 (RISA-^{125}I); protease activity on hemoglobin will liberate tyrosine which can be estimated spectrophotometrically at 280 nm; with RISA, liberated tyrosine-^{125}I, as estimated by isotopic procedures, is an index of proteolytic activity.

Pepsin activity can be distinguished from the total protease activity by estimation of the 3,5-diiodotyrosine liberated from *N*-acetyl-L-phenylalanyl-3,5-diiodotyrosine substrate at pH 2.1. Pepsin will react on this substrate; gastricsin will not. Normal gastric juice protease activity ranges from 200 to 1200 μg total protease activity/mL and 50 to 300 μg pepsin/mL. The presence of bile, blood, saliva or excess mucus in the sample will decrease both acidity and gastric protease activity.

Gastrin, cholecystokinin, secretin and pancreozymin are gastrointestinal hormones.[24] The role of gastrin and its interaction with other gastrointestinal hormones in the etiology and proliferation of ulcer disease is of recent interest. Accurate RIA techniques have been developed for gastrin and secretin-6-tyrosine due to the availability of a pure synthetic polypeptide. Biological assays based on the effect of

these substances on gastric, pancreatic and biliary secretion also have been used.

Gastrin is found in various species in two forms, G-I and G-II. The only difference is in sulfation of the 12-tyrosyl residue in G-II of the heptadecapeptide amides. Gastrin is found primarily in the gastrin-producing cells (G-cells) of the antral mucosa. The C-terminal tetrapeptide represents the biologically active part of the molecule. Gastrin infusion will stimulate secretion of gastric acid, pepsin and intrinsic factor. It has a slight secretin-like effect and a powerful pancreozymin-like effect on pancreatic secretion. Gastrin also stimulates bile flow. The instillation of HCl into the stomach will inhibit gastrin release; protein and meal stimulation will increase serum gastrin.

The RIA of serum gastrin is of diagnostic value in the ZE syndrome, pernicious anemia and duodenal ulcer. Basal serum gastrin levels in the normal individual are 20 to 30 μg/mL and increase about 2-fold after a protein meal stimulus. Basal serum gastrin levels in duodenal ulcer are normal or slightly elevated, but increase 4- to 5-fold after a protein-meal stimulus. Basal serum gastrin levels are elevated in ZE to 500 to 4000 pg/mL due to the presence of a gastrin-producing tumor. The ZE patient is uniquely sensitive to IV calcium stimulation which will increase both gastric acid secretion and serum gastrin in this syndrome. Basal serum gastrin levels also are elevated in gastric hyposecretion as seen in pernicious anemia and Type A gastritis, and in chronic renal failure due to the decreased metabolic turnover of gastrin in the kidney.

The RIA of serum gastrin is based on the competition of gastrin in test sample with ^{125}I-gastrin for gastrin antibody binding sites. The antibodies used in this procedure are usually cospecific for G-I and G-II. However, they detect all forms of circulating gastrin, ie, Big-Big Gastrin (G-39), Big Gastrin (mol wt 7000; G-33), gastrin heptadecapeptide (G-17, mol wt 2200), G-13 and G-8 (mini-gastrin). The Big components can be converted to gastrin by trypsin hydrolysis. The significance of changes in the ratio of the circulating gastrins is not known, but it has been suggested that G-39 and G-33 predominate in the basal state and cleave to G-17, which is the major serum form after a protein meal.

Other Body Fluids

Physical, chemical and microscopic examination of cerebrospinal fluid, synovial fluid, human milk, transudates and exudates also are performed by the clinical laboratory. The principles of the various determinations are similar to those described for blood and urine.

Microbiology

Clinical medical microbiology is a science which is concerned with the isolation and identification of disease-producing microorganisms, ie, bacteria, fungi (including yeast), viruses, rickettsia and parasites. The techniques employed in the isolation and identification of the suspect organism(s) involve the propagation on suitable primary culture media, selective isolation on special culture media, use of suitable living host material (mouse, embryonated egg, tissue culture, etc), determination of morphological and, where applicable, staining characteristics of the organism and confirmation by biochemical and/or immunochemical analysis. Suitable animal inoculation, where applicable, may be employed to determine pathogenicity. Site, timing, technique (aseptic), instrumentation, and transportation of clinical specimens (blood, urine, feces, cerebrospinal fluid, etc) are prime variables involved in the final differentiation and confirmation process.

Rapid manual enzymatic and immunological test kits have been introduced to identify pathogens for cerebrospinal fluid analysis. The latex-agglutination test coats a specific antibody onto latex particles and when an antigen is present, the latex particles are visible.[25] In the coagglutination test, the specific antibody is bound to protein A on the surface of a staphylococcal cell and the presence of antigen produces agglutination.[25]

Staphylococcus aureus (*Micrococcus pyogenes* var *aureus*) is a Gram-positive coccus frequently found on normal human skin and mucous membranes and frequently associated with abscesses, septicemia, endocarditis and osteomyelitis. Some strains elaborate an exotoxin capable of causing food poisoning. The primary isolation is on blood agar and in thioglycollate broth. With feces and other heavily contaminated specimens, phenylethyl alcohol agar and/or mannitol-salt agar should be inoculated to suppress growth of other bacteria. The identification of pathogenic staphylococci is based on colonial (pigmentation) and microscopic morphology (grape-like clusters), positive catalase production, positive coagulase production (staphylocoagulase-plasma clotting factor) and positive mannitol fermentation.

Streptococcus pyogenes is another Gram-positive coccus frequently associated with tonsillitis or pharyngitis, erysipelas, pyoderma and endocarditis. Neopeptone agar containing 5% defibrinated sheep blood is preferred for primary isolation and to demonstrate characteristic hemolysin production by observing a zone of clear (beta) hemolysis around the colonies on blood agar. Streptococcal groups are identified by precipitin tests with group-specific antisera for A, B, C, D, F and G. Streptex (*Wellcome Diagnostics*) uses a latex agglutination system for identifying the Lancefield group of streptococci. Other groups usually are not associated with human clinical materials.

Legionella pneumophilia identification includes specimen cultures on lung tissue or sterile body fluids (eg, pleural fluid or pericardial fluid). Direct fluorescent antibody method is a test for *L pneumophilia*. Organisms are best seen in the acute stage of the disease. Since the antiserum is species-specific, polyvalent antisera are necessary for identification.

Neisseria gonorrheae is a Gram-negative diplococcus associated with the venereal disease gonorrhea. The identification is based on the primary isolation of the gonococcus from urethral exudates on chocolate agar or Thayer-Martin (TM) medium. The microscopic observation of Gram-negative intracellular diplococci resembling the gonococcus constitutes a presumptively positive diagnosis of gonorrhea. Confirmation of the oxidase enzyme activity of the gonococci is performed by a reaction with *p*-dimethylaminoaniline which turns oxidase-positive colonies black. A positive oxidase test by Gram-negative diplococci isolated on TM medium constitutes a presumptively positive test for *N gonorrheae*. Final identification rests on typical sugar fermentation or specific (fluorescent antibody) staining.

Neisseria meningitidis is the primary cause of bacterial meningitis and septicemia. The primary isolation is based on culturing of a specimen (blood, spinal fluid or nasopharyngeal secretions) on a Mueller-Hinton medium or chocolate agar containing a vancomycin-colistimethate-nystatin antibiotic mixture. The confirmation of the isolate by biochemical reactions (positive oxidase, positive catalase, etc) and serological agglutination with group-specific (A, B and C) antiserum is used in the differentiation. Young cultures of groups A and C may show capsular swelling (Quellung reaction) in the presence of a specific antiserum.

The enteric bacilli (*Enterobacteriaceae*) are Gram-negative, nonsporulating rods associated with dysentery (*Shigella sp*), typhoid fever (*Salmonella typhi*), urinary tract and tissue infections (*Escherichia coli, Proteus sp* and *Pseudomonas sp*), and pulmonary infections (*Klebsiella sp*). The primary isolation of enteric bacilli is on selective and differential infusion agar such as MacConkey and eosin-methyl-

ene blue (EMB), and enrichment media such as selenite broth and tetrathionate broth. The primary isolation of *Salmonella sp.* is on Leifson's deoxycholate citrate agar (LDC) or *Salmonella-shigella* agar (SS); if *Salmonella typhi* is suspected, brilliant green agar (BG) and bismuth sulfite agar (BS) may be used and would constitute a presumptively positive diagnosis of *S typhi*.

The confirmation and identification of enteric bacilli may be performed by serological tests and biochemical reactions: H_2S production (triple-sugar iron agar), indole production, acetylmethylcarbinol production, citrate utilization, urease, lysine and arginine decarboxylase and phenylalanine deaminase activity. Enterotube (*Roche Diagnostics*) employs conventional media to perform 11 standard biochemical tests which can be inoculated simultaneously in one compartmented tube, with a single bacterial colony. The serological identification of *Salmonella* and *Shigella sp* is based on the agglutination of antigens that fall into three categories: "K" capsular (*Klebsiella sp* and *Shigella sp*), "O" (*Salmonella sp, Arizona sp, E coli, Shigella sp*, etc) and "H" flagellar (*Salmonella sp*).

Other Gram-negative rods of medical importance are the hemophilic bacilli (*Bordetella pertussis*, whooping cough and *Hemophilus influenzae*, bacterial meningitis), the hemorrhagic bacilli (*Pasteurella pestis*, bubonic plague, and *P tularensis*, tularemia) and pyrogenic bacillus (*Brucella melitensis*, undulant fever).

Spore-forming Gram-positive rods of medical importance belong to the genus *Clostridium*, which are associated with tetanus (*Cl tetani*), gas gangrene (*Cl perfringens* or *welchii*) and botulism (*Cl botulinum*). The isolation of these organisms requires anaerobic conditions. Once the strain to be identified is obtained in pure culture by single-colony selection, its morphological characteristics are noted; the strain then is grown in a variety of definitive media to determine catalase activity, hydrogen peroxide decomposition and fermentation or hydrolysis of carbohydrates and organic acids. The analysis of fermentation products (gas chromatography) also is used for the identification of pathogenic anaerobic *Clostridia*. The major clostridial exotoxin type can be determined by typing with specific antitoxin sera. A Gram-positive, aerobic, spore-former of medical importance is *Bacillus anthracis*, responsible for anthrax, a disease of animals transmissible to man.

The mycobacteria are acid-fast bacilli associated with tuberculosis in man (*Mycobacterium tuberculosis*), in cattle (*Mycobacterium bovis*) and leprosy (*Mycobacterium leprae*). Tubercle bacilli in man are isolated from sputum cultured on a tubed or bottled egg medium (Lowenstein-Jensen) following enzymatic digestion and concentration of the specimens. A provisional diagnosis of tuberculosis usually is made by demonstrating acid-fast bacilli microscopically, X-ray diagnosis and a positive tuberculin skin test.

Other weakly and partially acid-fast bacilli of medical importance are members of the *Actinomycetales*, *Nocardia asteroides* and *Nocardia brasiliensis*, which are responsible for severe pulmonary infections and cutaneous and subcutaneous abscesses.

Bacteriophages (phages) are a special group of viruses that are hosted by bacteria. Any given phage is highly host-specific and when in contact, lysis of the host occurs (phage-typing). They are used primarily as epidemiological tools in subtyping strains of *E coli*, staphylococci or *Salmonella sp* that are presumed to be related epidemiologically. Phages also furnish ideal material for studying host-parasite relationships and virus multiplication.

The medically important fungal diseases include the superficial mycoses, ie, fungal invasion is restricted to the outermost layers of the skin or to the hair shafts (*Microsporum audouini*, ringworm of the scalp, *Trichophyton sp*, athlete's foot and *Epidermophyton floccosum*, *Tinea pedis*) and the systemic pathogenic fungi (*Blastomyces dermatitidis, Coccidioides immitis, Histoplasma capsulatum, Candida albicans*). The diagnosis of the causative agent is based on the isolation of organisms on Sabouraud's dextrose agar or trypticase soy agar with or without cycloheximide and chloramphenicol to suppress the growth of saprophytic fungi and bacteria, macroscopic examination of morphological characteristics and microscopic examination using KOH or lactophenol cotton-blue stain. Biochemical reactions usually are limited to *Candida sp*. Immunological reactions include skin tests, where applicable, agglutination tests, such as latex particle agglutination for histoplasmosis and tube precipitin and complement-fixation tests.

An *antimicrobial susceptibility test* is a determination of the least amount of an antimicrobial chemotherapeutic agent that will inhibit the growth of a microorganism *in vitro*, using a tube-dilution method, agar cup or disk-diffusion method. The test may function as an aid in the selection of a chemotherapeutic agent by the physician. Also, the concentration of antimicrobial agents in body fluids may be determined by biological assay with an organism of known susceptibility for the specific agent.

The laboratory diagnosis of *viral infections* is based upon (1) examination of the infected tissues for pathognomonic changes or for the presence of viral material; (2) isolation and identification of the viral agent; (3) demonstration of a significant increase in antibody titer to a given virus during the course of the illness; (4) detection of viral antigens in lesions—using fluorescein-labeled antibodies and (5) electron microscopic examination of vesicular fluids or tissue extracts. Blood is used for serological tests but seldom for virus isolation. Acute and convalescent-phase blood specimens must be examined in parallel to determine whether or not antibodies have appeared or increased in titer during the course of the disease. Some examples of human viral infections are respiratory infections (Adenovirus group); diseases of the nervous system, ie, polio and coxsackie viruses of the picornavirus group; smallpox (poxvirus group); measles (paramyxovirus group); chicken pox (herpesvirus group) and influenza (myxovirus group).

Members of *Mycoplasmatacea* pleuropneumonia-like organisms (PPLO) are of a range of size similar to the larger viruses. They are highly pleomorphic because they lack a rigid cell wall, they can reproduce in cell-free media and they do not revert to or from bacterial parental forms as the L-forms. Specimens (sputum, bronchial secretions, urinary sediment, etc) for the primary isolation of mycoplasmas (*M pneumoniae, M hominis*, etc) should be cultured on agar media containing peptone, serum, ascitic fluid, whole blood or egg yolk. The species identification may be by growth inhibition on agar medium containing type-specific rabbit antisera. Antigenic variants or subspecies may be detected by immunodiffusion. Various PPLO are pathogenic, parasitic or saprophytic. Mycoplasmas have a predilection for mucous membranes and are associated with primary atypical pneumonia and bronchitis.

Clinical parasitology is a science which is concerned with the parasitic protozoa (amoeba), the helminths (cestodes, tapeworms; trematodes, flukes; nematodes, roundworms) and the arthropods. The identification of protozoan ova is based on detailed microscopic morphology (nuclei, etc) using wet mounts (saline or iodine) or stained preparations (iron hematoxylin, etc) obtained from fecal specimens (fresh or preserved with polyvinyl alcohol), which are concentrated by sedimentation, centrifugation or flotation techniques. Trophozoite and/or cystic stages may be detected in fecal specimens associated with intestinal protozoa as in amoebic dysentery caused by *Entamoeba histolytica*.

The commonly encountered helminths are *Necator americanus* (hookworm), *Trichuris trichiura* (whipworm) and *Enterobius vermicularis* (pinworm); they are identified by

characteristic ova. Characterization of tapeworm segments (proglottids) or head (scolex) in a fecal specimen will differentiate *Taenia saginata* (beef tapeworm) from *Taenia solium* (pork tapeworm). Eggs of *T solium* and *T saginata* cannot be differentiated on a morphological basis.

Adult flukes oviposit a characteristic egg which may reach the urine, sputum or feces. *Schistosoma japonicum* eggs have a small, indistinct spine; *S mansoni*, a distinct, large, lateral spine; and *S haematobium*, a distinct terminal spine.

Arthropoda constitute the largest of the animal phyla which are characterized by a segmented body with the segments usually grouped in two or three distinct body regions, by a chitinous exoskeleton, several pairs of jointed appendages and characteristic internal organs. Most arthropods can be preserved in 70% alcohol. They are of medical importance since they can infest man and cause mechanical trauma or produce hypersensitivity from repeated exposure (*Cimex lectularius*, the bedbug) or by toxin injection (*Latrodectus mactans*, the black widow spider), by skin invasion (*Sarcoptes scabiei*, the itch mite) and by transmitting disease (*Anopheles* mosquitoes, malaria), and *Yersinia pestis* in fleas (plague).

The serodiagnosis of parasitic diseases includes the following immunodiagnostic tests: complement-fixation (trichinosis), precipitin test (schistosomiasis), bentonite flocculation (ascariasis), hemagglutination (echinococcosis), latex agglutination (trichinosis), cholesterol flocculation (schistosomiasis), fluorescent antibody (malaria) and methylene blue dye test (toxoplasmosis).

Immunochemistry

Clinical immunopathology[26] includes *general immunology* (immunofluorescence, immunodiffusion, immunoelectrophoresis and agglutination tests), *radioimmunoassay* (RIA-hormones, vitamins, drugs, immunoglobulins), *tissue typing* (histocompatibility tests in organ transplants), *cellular immunology*, *cancer immunology* and *immunohematology*. Examples of each of these disciplines are discussed in this section and other parts of this chapter.

The ELISA, *enzyme-linked immunosorbent assay*, detects antibodies by an indirect technique using enzyme-linked antibodies to label antigenic substances in tissue or body fluid. The antigen is attached to a solid matrix and reacts with a specimen that may contain a complimentary antibody. The antihuman globulin, which is conjugated with the enzyme, is added and the antigen reacts with the bound antibody of the patient. By adding the substrate molecule the enzyme is detected. This system has been used to identify antibodies to viruses, parasites, bacterial products and quantitation of some drugs.

Antibody response is a complex process involving the lymphoid cell system response to foreign stimulus or antigen. Hematopoietic cells in the fetal yolk sac, liver or marrow develop into lymphoid stem cells which, in turn, differentiate into T-lymphocytes of thymic origin and B-lymphocytes of bone-marrow origin. The T-cells further differentiate into lymphoblasts which are responsible for *cell-mediated cellular immunity* (graft vs host reaction, tissue transplant rejection, tuberculin skin testing, *delayed-type hypersensitivity*). B-cells differentiate into plasma cells which are responsible for humoral immunity which is mediated by circulating serum immunoglobulins (*immediate-type hypersensitivity*).

Macrophages can cooperate in presentation of antigen to the T- or B-lymphoblasts. Cooperation between T- and B-cells, immunological memory, development of immune tolerance to antigens and genetic control of the immune response are integral properties of the immune system and are related to development of immune deficiency and autoimmune disease.

The identification and determination of *immunoglobulins* (IgG, IgM, IgA) by radial immunodiffusion and immunoelectrophoresis have been discussed under *Proteins*. *IgM* (γM) is the earliest antibody found in the primary immune response and falls rapidly after the onset of IgG antibody synthesis. *IgG* (γG) is the major class of antibody in both the primary and secondary immune response. IgG can cross the placenta to provide the early forms of antibody protection for the newborn. IgG and IgM can participate in the complement fixation reaction. *IgA* (γA) is found predominantly in saliva and secretions of the gastrointestinal and respiratory tracts. In contrast to IgM and IgG, only a small portion of total IgA is found in blood. IgA functions in protection against pathogens that enter the host through the respiratory or gastrointestinal tract. *IgD* (γD) is found in trace quantities in sera and its function is unknown. *IgE* (γE) is probably the most important antibody in acute hypersensitivity or allergic reactions. Reaction of mast cell- or basophil-bound IgE with antigen initiates the release of histamine, slow-reacting substance (SRS), serotonin and bradykinin and the subsequent allergic response. IgE is best quantitated by RIA. Mean serum levels (mg%) in healthy adults are IgG 1200 ± 500, IgA 210 ± 140, IgM 140 ± 70, IgD 3 and IgE <0.1.

Heterophile antibodies are agglutinins which are capable of reacting with antigens that are entirely unrelated to those which stimulate their production. These antibodies, which occur in the serum of patients with infectious mononucleosis or serum sickness, will agglutinate formalized horse erythrocytes. In order to distinguish the specific *heterophile agglutinins of infectious mononucleosis*, the serum sample is mixed with guinea-pig kidney tissue or beef erythrocyte stromata; the infectious mononucleosis antibody will be absorbed and inactivated by the beef cells but not by the kidney tissue, and subsequent agglutination of horse erythrocytes will occur only in the kidney-tissue system. This test is used to detect infectious mononucleosis even prior to clinical symptoms. The heterophile titer has no relation to the course or severity of the disease.

Two protein constituents of human plasma, *rheumatoid factor* (RF) and *C-reactive protein* (CRP) are of value in the differential diagnosis of rheumatoid diseases. CRP is a protein present in the serum of patients in the acute stages of bacterial and viral infections, collagen diseases and other inflammatory processes. The presence of this antigen in serum is detected by agglutination of polystyrene latex particles sensitized with specific CRP antibody globulin. In the management of rheumatic fever, decreases in CRP blood levels are used to measure the effectiveness of therapy.

Rheumatoid arthritis is characterized by the presence of a reactive group of macroglobulins known as RF in blood and synovial fluid. RF is a protein of the IgM globulin fraction and is regarded as an autoantibody against antigenic determinants of IgG. Analysis of RF is based on agglutination procedures employing polystyrene latex particles coated with a layer of adsorbed human gamma globulin. The RF-antibody reaction causes a visible agglutination of the inert latex particles. CRP is not elevated in rheumatoid arthritis.

β-Hemolytic streptococci, the causative agent in rheumatic fever, produce streptolysin O and S, streptokinase, hyaluronidase, desoxyribonuclease and NADase in the body. The growth of streptococci in tissue with elaboration of these proteins serves as the antigenic stimulus to evoke the production of specific antibodies (eg, *antistreptolysin-O*, *ASO*). The quantitation of the antibody titer to these enzymes is an index of the strength of the antigenic stimulus and the extent of the streptococcal infection. These antibodies can be detected by latex agglutination (ASO) or tests

dependent on the inhibition of enzyme action by the antibody (anti-hyaluronidase inhibition of hyaluronic acid depolymerization by hyaluronidase).

The laboratory diagnosis of *syphilis* (treponemal disease) and the evaluation of a chemotherapeutic approach is based on serological tests. Demonstration of an antibody-like substance, *reagin*, or of true antitreponemal antibody in the serum of infected individuals is accomplished by complement fixation or flocculation tests for reagin, or immunofluorescent techniques for treponemal antibody.

In the *complement fixation* tests (Kolmer CF), reagin reacts with a complex phosphatidic acid antigen (cardiolipin) and complement; the complement is bound and will not lyse hemolysin-sensitized red cells which were added in the second phase of the test. In normal serum the reagin-cardiolipin complex is not formed and the complement is free to react with hemolysin and lyse the erythrocytes.

Flocculation tests for determining of syphilis use a cardiolipin-lecithin-cholesterol antigen which clumps in the presence of serum reagin occurring in nontreponemal diseases and syphilis (*Venereal Disease Research Laboratory—VDRL Test; rapid plasma reagin—RPR test*).

Treponemal antibody can be detected also by the reaction of the patient's serum with treponemal antigen and subsequent confirmation with fluorescein-labeled antihuman globulin as an indicator of primary antigen-antibody reaction (*fluorescent treponemal antibody-FTA test*). The patient's serum can be treated with an extract of treponemes prior to the FTA test to remove interfering antibodies and eliminate biological false-positives (FTA-Abs Test). False-positives occur in related treponematosis such as yaws, pinta and bejel. Increased reagin titers also occur in malaria, leprosy, infectious mononucleosis, chronic rheumatoid arthritis or systemic lupus erythematosus and in patients on hydralazine therapy.

Febrile antibodies are present in the serum of patients with certain bacterial or rickettsial infections (spotted, typhus or Q fever). In typhus the patient's serum contains a febrile antibody which will agglutinate a suspension of *Proteus OX-19* bacteria (Weil-Felix Reaction). *Salmonella* O-H, *Pasteurella tularensis* and *Brucella abortus* antigens are used in febrile antibody tests for diagnosis of typhoid or paratyphoid fever, tularemia and brucellosis, respectively.

Toxoplasmosis is a major cause of birth defects. An expectant mother may become infected with oocysts in uncooked meat, or from cat fur, and infect the fetus transplacentally. Toxoplasmosis testing is based on detecting serum antibody by a hemagglutination procedure. Red cells sensitized by exposure to toxoplasmosis antigen are agglutinated by the specific antibody.

Radioimmunoassay (RIA)[6,27] has been mentioned in various sections of this chapter as an analytical tool in the measurement of hormones, immunoglobulins, drugs and steroids. The basic principle of RIA is

$$Ag^* + Ag + Ab \rightleftharpoons Ag^*Ab + AgAb + Ag^* + Ag \qquad (4)$$

RIA is not to be confused with the *specific reactor assay* using labeled antigen and nonantibody protein receptors which is used for vitamin B_{12}, T^4, T^3 and cortisol assays.

All procedures are based on the observation that radiolabeled antigens (Ag*) compete with nonlabeled antigen (Ag) for binding sites on specific antibody (Ab) in the formation of antigen-antibody complexes (Ag*Ab, AgAb). When increasing amounts of Ag are added to the assay, the binding sites of Ab are saturated progressively and the antibody can bind less Ag*. Therefore, the ratio of bound to free Ag* (B/F) or % Ag* bound is a direct index of the concentration of Ag in the assay.

The requirements for RIA are (1) preparation and characterization of Ag (2) radiolabeling of Ag, (3) preparation of specific Ab and (4) development of the assay system and methods to separate free (Ag, Ag*) from antibody bound (AgAb, Ag*Ab) antigen.

Antigens can be prepared from natural tissue sources or preferably synthesized. ^3H, ^{14}C or ^{125}I-labeled antigens are used routinely in the assay. The biological and immunochemical activity of the antigen must not be altered in the tagging procedure, and the specific activity of Ag* must be extremely high so that tracer quantities can be used in the assay. Tritium labeling and iodination (^{125}I) produce the highest specific activity, but also increase susceptibility of Ag* to internal degradation and self-radiolysis, in contrast to ^{14}C. In many instances, the original antigen cannot be iodinated, but can be altered chemically in such a way as to retain full antigenic cross-reactivity in RIA; eg, cyclic AMP has no tyrosyl or histidyl residue for iodination; ^{125}I-succinylcyclic AMP-tyrosine methyl ester retains full cross-reactivity with antibodies to cyclic AMP and is used in the assay.

Hormones, steroids and drug substances are *haptens*. They do not produce the antibody response when injected by themselves, but will produce antibodies specific for the hapten when injected as a hapten-protein carrier conjugate. Gastrin (hapten) is coupled to albumin (protein-carrier) by treatment with carbodiimides (CCD), which couple functional carboxyl, amino, alcohol, phosphate or thiol groups. Morphine must be converted to the 3-*O*-carboxymethyl derivative prior to CCD coupling with albumin to provide a functional coupling group in the hapten. The hapten-conjugate usually is emulsified in a mineral oil preparation of killed *Mycobacterium* (Complete Freunds Adjuvant) and injected intradermally in rabbits or guinea pigs on several occasions. The serum antibody must have both high specificity and affinity for the antigens.

The *assay system* contains Ag*, sample-containing endogenous Ag or a standard Ag and antibody, at specified pH (6.5 to 8.5). After incubation at 5 to 37° for anywhere from 1 hour to several days, free and antibody-bound antigen must be separated. This is accomplished by *double-antibody technique, solid-phase RIA, resin techniques* or *salt or solvent precipitation*. In the double-antibody technique, antiglobulin (Ab′) serum is added to the assay system after incubation. Ab-Ag* and Ab-Ag complexes are antibody-globulin antigen complexes. The antiglobulin will react to form insoluble Ab′-Ab-Ag* and Ab′-Ab-Ag complexes, which can be removed by centrifugation. The free Ag*, Ag is in the supernate.

The solid phase RIA is performed by coating tubes with Ab; Ag and Ag* react, compete and bind with Ab on the wall of tube. Unreacted Ag and Ag* is separated by decanting and rinsing the tube. Ab also can be bound covalently with isothiocyanate to dextran gel particles. Ag and Ag* will compete and bind with Ab on particles. Bound antigen then can be separated from free antigen by centrifugation.

RIA has been applied to analysis of hormones (ACTH, angiotensin I and II, gastrin, HCG, FSH, GH, glucagon, HLH, HPL, insulin, thyroxine), steroid hormones (aldosterone, androstenedione, glucocorticoids, testosterone, estrones, progesterone), drug substances (digoxin, digitoxin, amphetamines, barbiturates, morphine, LSD, ouabain), endogenous substances (cyclic AMP, cyclic GMP, prostaglandins, immunoglobulins, hepatitis antigen, carcinoembryonic antigen—CEA). Examples of the specific assays are discussed in other sections.

CEA and AFP (α-1-fetoprotein) are proteins found in fetal tissue. CEA analysis was first proposed as a specific test for the early detection of bowel cancer. Although the test does not have absolute specificity for this disease, it may prove of value as a diagnostic aid and therapy monitor. CEA can be detected by RIA. Serum levels >2.5 ng CEA/mL are found in 60 to 70% of patients with adenocarci-

noma of the colon; positive levels also are found in lower percentages in carcinomas of the pancreas, stomach, liver, breast, endometrium, ovary, kidney and bronchus, as well as in other conditions such as gastrointestinal polyps, colitis, diverticulitis and cirrhosis. CEA appears to be associated primarily with tumors of entodermally derived epithelial tissue. The similarity between CEA and cell-surface glycoproteins and sialic acids has stimulated considerable research interest in a new approach to cancer chemotherapy.

The study of *tissue-transplantation antigens* is an important factor in studies on tissue and organ transplants. ABO blood group antigens are involved in survival of skin and renal grafts. Because of the presence of natural occurring anti- A and B, avoidance of ABO incompatibility is important in clinical grafting. The *HL-A antigens* are found on tissue and on the white cells. There is one major histocompatibility locus, comprising a number of alleles or linked genes, on a single chromosome segment. Each allele controls four to five groups of major transplantation antigens. These HL-A isoantigens affect the survival of allogenic tissue grafts and organ transplants. HL-A antigens can be typed by a leukoagglutination method in which the patient's or donor's white cells are reacted with specific HL-A antisera. HL-A typing also can be performed by a cytotoxicity test in which lymphocytes are mixed with antisera and complement. The antibody can destroy the lymphocytes if a corresponding antigen is present on the cell surface.

References

1. Mitruka BM, Rawnsley AM: *Clinical and Hematological Reference Values in Normal Experimental Animals and Normal Humans*, Yearbook Medical Pub, Chicago, 1981.
2. Christensen RL, Triplett DA: *Lab Med, 13(11):* 666-672, 1982.
3. Bollinger P, Brailas CD, Drewinko B: Evaluation of whole-blood platelet analyzers. *Ibid 14:* 492, 1983.
4. *Central File for Rare Donors*, Am Assoc Blood Banks, Milwaukee.
5. *ABO and Rh Systems*, Ortho Diagnostics, Raritan NJ 1969.
6. Berson S, Yalow R: *Gastroenterol 62:* 1061, 1972.
7. *Fed Reg 37FR17419*, Aug 26, 1972.
8. Broughton PMG, Dawson JB: *Advan Clin Chem 15:* 288, 1972.
9. Mears T, Young D: *Am J Clin Pathol 50:* 411, 1968.
10. Meinke W: *Anal Chem 43:* 28A, 1971.
11. Dybaker R: *Std Methods Clin Chem 6:* 223, 1970.
12. Flokstra JH, Soda JA: *Am J Med Sci 254:* 429, 1967.
13. Young DS *et al: Clin Chem 21:* 1D, 1975.
14. Constantino NV, Kabat HF: *Am J Hosp Pharm 30:* 24, 1973.
15. Peterson CM: *Diagn Med:* 78, Jul/Aug 1980.
16. *Radial Immunodiffusion and Immunoelectrophoreses for Qualitation and Quantitation of Immunoglobulins* (DHEW Publ HSM-72-8102), USD-HEW, Washington DC, 1972.
17. *Bull WHO 43:* 891, 1970.
18. Statland BE: *Clinical Decision Levels For Lab Tests*, Med Econ Co Inc., Oradell NJ, 1983.
19. White W *et al: Practical Automation for the Clinical Laboratory*, Mosby, St Louis, 1972.
20. Szilagyi G, Aning V, Karmen A: *J Clin Lab Automation 3:* 117, 1983.
21. Bradley GM: Fecal analysis. *Diagn Med* 63 Mar/Apr 1980.
22. Rubenstein K *et al: Biochem Biophys Res Comm 47:* 846, 1972.
23. Baron J: *Scand J Gastroenterol 5:* 9, 1970.
24. Jorpes J, Mott V: *Secretin, CCK, Pancreozymin and Gastrin*, Springer Verlag, New York, 1973.
25. Kuhn PJ: *Mod Lab Observer* 108, Sept 1983.
26. Feldman M, Nossal GJ: *Quart Rev Biol 47:* 269, 1972.
27. Skelley DS, *et al: Clin Chem 19:* 146, 1973.

Bibliography

Wintrobe M: *Clinical Hematology*, 6th ed, Lea & Febiger, Philadelphia, 1967.

Lynch MJ: *Medical Laboratory Technology*, 2nd ed, Saunders, Philadelphia, 1969.

Faulkner W, King J: *Manual Clinical Laboratory Procedures*, Chemical Rubber Co, Cleveland, 1970.

Faulkner W *et al: Handbook Clinical Laboratory Data*, Chemical Rubber Co, Cleveland, 1980.

Roth K, Saunders A: *Evaluation of Methods for White Cell Identification and Counting-Advances in Automated Analysis*, Technicon Intl Congr, 1970.

Dacie J, Lewis S: *Practical Hematology*, 3rd ed, Churchill, London, 1963.

Frankel S, Reitman S, eds: *Clinical Laboratory Methods and Diagnosis*, 6th ed, Mosby, St. Louis, 1963.

Wintrobe MM: *Laboratory Medicine-Hematology*, 2nd ed, Mosby, St. Louis, 1962.

Manual of Blood Coagulation Technics, 2nd ed, Warner-Chilcott, Morris Plans NJ, 1966.

Detection of Fibrinogen Degradation Products, Wellcome Res Labs, England, 1973.

A Manual of Methods for the Coagulation Laboratory, BD & Co, Rutherford NJ, 1965.

Technical Methods and Procedures of the American Association of Blood Banks, Am Assoc Blood Banks, Chicago, 1962.

Standards for Blood Transfusion Service, 4th ed, Am Assoc Blood Banks, Chicago, 1963.

Griffiths JJ, Elliott J: *Blood Bank Procedures*, Dade Reagents, Miami, 1967.

Chromatography in Mass Screening for Disorders of Amino Acid Metabolism, Hyland, Los Angeles, 1966.

Rosalki S, Wilkinson J: *Diagnostic Enzymology*, Dade Reagents, Miami, 1966.

Wilkinson J: *Introduction to Diagnostic Enzymology*, Edward Arnold, Ltd. London, 1962.

Davidsohn I, Henry J: *Todd-Sanford Clinical Diagnosis by Laboratory Methods*, 15th ed, Saunders, Philadelphia, 1974.

Peron FG, Caldwell BV: *Immunologic Methods in Steroid Determination*, Appleton-Century, New York, 1970.

Winsten S, Dalal F: *Clinical Laboratory Procedures for Nonroutine Problems*, Chemical Rubber Co, Cleveland, 1972.

Specialized Diagnostic Laboratory Tests, Bioscience Labs, Van Nuys CA, 1971.

Kark RM *et al: A Primer of Urinalysis*, 2nd ed, Harper & Row, New York, 1963.

Sunderman FW, Sunderman FW, Jr: *Laboratory Diagnosis of Renal Diseases*, Warren H Green, Inc, St Louis, 1970.

Faust E, Russell P: *Clinical Parasitology*, 7th ed, Lea & Febiger, Philadelphia, 1964.

Sunshine I: *Manual of Analytical Toxicology*, Chemical Rubber Co., Cleveland, 1972.

Clarke, E: *Isolation and Identification of Drugs*, vols 1 and 2, Pharmaceutical Press, London, 1969 and 1975.

Blair JE *et al: Manual of Clinical Microbiology*, Williams & Wilkins, Baltimore, 1970.

Edwards PR, Ewing WH: *Identification of Enterobacteriaceae*, 3rd ed, Burgess, Minneapolis, 1972.

Holdeman LV, Moore WEC: *Anaerobe Laboratory Manual*, 2nd ed, VPI Anaerobe Lab, Blacksburg VA, 1973.

Connant NF *et al: Manual of Clinical Mycology*, 3rd ed, Saunders, Philadelphia, 1971.

Bach F, Good R: *Clinical Immunobiology*, Academic, New York, 1972.

Clinical RIA. *Lab Manag:* May 1973.

Manual of tissue typing techniques. *Natl Inst All Infect Dis Bull:* 1972.

Directory of Rare Analysis. *Clin Chem 23:* 323, 1977.

Doucet LD, *Medical Technology Review*, Lippincott, Philadelphia, 1981.

Hansten PD, *Drug Interactions*, 3rd ed, Lea & Febiger, Philadelphia, 1975.

Kaplan A, Szabo LL, *Clinical Chemistry: Interpretation and Techniques*, 2nd ed, Lea & Febiger, Philadelphia, 1983.

Miller SE, Weller JM, *Textbook of Clinical Pathology*, 8th ed, Williams & Wilkins, Baltimore, 1971.

Peacock J, Tomar R, *Manual of Laboratory Immunology*, Lea & Febiger, Philadelphia, 1980.

Pertinent Reference Journals

Advan Clin Chem	*J Clin Lab Automation*
Am J Clin Pathol	*Diagn Med*
Am Clin Prod Rev	*J Lab Clin Med*
Am J Hosp Pharm	*Lab Med*
Am J Med Technol	*Lab Notes Med Diag*
Anal Chem	*Med Lab Obs*
BioTechniques	*Med Lab Tech*
Clin Chem	*Scand J Clin Lab Invest*
Clin Chim Acta	*Std Methods Clin Chem*

CHAPTER 35

Drug Absorption, Action and Disposition

Stewart C Harvey, PhD

Professor of Pharmacology
School of Medicine, University of Utah
Salt Lake City, UT 84132

Although drugs differ widely in their pharmacodynamic effects and clinical application, in penetrance, absorption and usual route of administration, in distribution among the body tissues and in disposition and mode of termination of action, there are certain general principles that help explain these differences. These principles have both pharmaceutic and therapeutic implications. They facilitate an understanding of both the features that are common to a class of drugs and the differentia among the members of that class.

In order for a drug to act it must be absorbed, transported to the appropriate tissue or organ, penetrate to the responding subcellular structure and elicit a response or alter ongoing processes. The drug may be distributed simultaneously or sequentially to a number of tissues, bound or stored, metabolized to inactive or active products or excreted. The history of a drug in the body is summarized in Fig 35-1. Each of the processes or events depicted relates importantly to therapeutic and toxic effects of a drug and to the mode of administration, and drug design must take each into account. Since the effect elicited hy a drug is its *raison d'etre*, *drug action* and *effect* will be discussed first in the text that follows, even though they are preceded by other events.

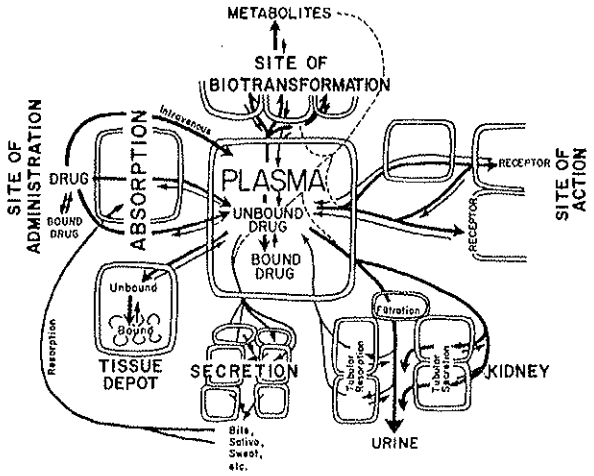


Fig 35-1. The absorption, distribution, action and elimination of a drug (arrows represent drug movement). Intravenous administration is the only process whereby a drug may enter a compartment without passing through a biological membrane. Note that drugs excreted in bile and saliva may be resorbed.

Drug Action and Effect

The word *drug* imposes an action–effect context within which the properties of a substance are described. The description of necessity must include the pertinent properties of the recipient of the drug. Thus, when a drug is defined as an analgesic, it is implied that the recipient reacts in a certain way, called pain,* to a noxious stimulus. Both because the pertinent properties are locked into the complex and somewhat imprecise biological context and because the types of possible response are many, descriptions of the properties of drugs tend to emphasize the qualitative features of the effects they elicit. Thus, a drug may be described as having analgesic, vasodepressor, convulsant, antibacterial, etc, properties. The specific effect (or use) categories into which the many drugs may be placed are the subject of Chapters 38 through 65 and will not be elaborated upon in this chapter. However, the description of a drug does not end with the enumeration of the responses it may elicit. There are certain intrinsic properties of the drug–recipient system that can be described in quantitative terms and which are essential to the full description of the drug and to the validation of the drug for specific uses. Under *Definitions and Concepts*, below, certain general terms are defined in qualitative language; under *Dose–Effect Relationships* the foundation is laid for an appreciation of some of the quantitative aspects of pharmacodynamics.

Definitions and Concepts

In the field of pharmacology, the vocabulary that is unique to the discipline is relatively small, and the general vocabulary is that of the biological sciences and chemistry. Nevertheless, there are a few definitions that are important to the proper understanding of pharmacology. It is necessary to differentiate among action, effect, selectivity, dose, potency and efficacy.

Action vs Effect—The *effect* of a drug is an *alteration of function* of the structure or process upon which the drug acts. It is common to use the term action as a synonym for effect. However, action precedes effect. *Action* is the *alteration of condition* that brings about the effect.

The final effect of a drug may be far removed from its site of action. For example, the diuresis subsequent to the ingestion of ethanol does not result from an action on the kidney but instead from a depression of activity in the supraopticohypophyseal region of the hypothalamus, which regulates the release of antidiuretic hormone from the posterior pituitary gland. The alteration of supraopticohypo-

* Sophisticated studies indicate that pain is not simply the *perception* of a certain kind of stimulus but, rather, a *reaction* to the perception of a variety of kinds of stimuli or stimulus patterns.

physeal function is, of course, also an effect of the drug, as is each subsequent change in the chain of events leading to diuresis. The action of ethanol was exerted only at the initial step, each subsequent effect being then the action to a following step.

Multiple Effects—No known drug is capable of exerting a single effect, although a number are known that appear to have a single mechanism of action. Multiple effects may derive from a single mechanism of action. For example, the inhibition of acetylcholinesterase by physostigmine will elicit an effect at every site where acetylcholine is produced, is potentially active, and is hydrolyzed by cholinesterase. Thus, physostigmine elicits a constellation of effects.

A drug also can cause multiple effects at several different sites by a single action at only one site, providing that the function initially altered at the site of action ramifies to control other functions at distant sites. Thus, a drug that suppresses steroid synthesis in the liver may not only lower serum cholesterol, impair nerve myelination and function and alter the condition of the skin (as a consequence of cholesterol deficiency) but also may affect digestive functions (because of a deficiency in bile acids) and alter adrenocortical and sexual hormonal balance.

Although a single action can give rise to multiple effects, most drugs exert multiple actions. The various actions may be related, as, for example, the sympathomimetic effects of metaraminol that accrue to its structural similarity to norepinephrine and its ability partially to suppress sympathetic responses because it occupies the catecholamine storage pools in lieu of norepinephrine; or the actions may be unrelated, as with the actions of morphine to interfere with the release of acetylcholine from certain autonomic nerves, block some actions of 5-hydroxytryptamine (serotonin) and release histamine. Many drugs bring about immunologic (allergic or hypersensitivity) responses that bear no relation to the other pharmacodynamic actions of the drug.

Selectivity—Despite the potential most drugs have for eliciting multiple effects, one effect is generally more readily elicitable than another. This differential responsiveness is called *selectivity*. It usually is considered to be a property of the drug, but it is also a property of the constitution and biodynamics of the recipient subject or patient.

Selectivity may come about in several ways. The subcellular structure (receptor) with which a drug combines to initiate one response may have a higher affinity for the drug than that for some other action. Atropine, for example, has a much higher affinity for muscarinic receptors (page 889) that subserve the function of sweating than it does for the nicotinic receptors (page 889) that subserve voluntary neuromuscular transmission, so that suppression of sweating can be achieved with only a tiny fraction of the dose necessary to cause paralysis of the skeletal muscles. A drug may be distributed unevenly, so that it reaches a higher concentration at one site than throughout the tissues generally; chloroquine is much more effective against hepatic than intestinal (colonic) amebiasis because it reaches a much higher concentration in the liver than in the wall of the colon. An affected function may be much more critical to or have less reserve in one organ than in another, so that a drug will be predisposed to elicit an effect at the more critical site. Some inhibitors of dopa decarboxylase (which is also 5-hydroxytryptophan decarboxylase) depress the synthesis of histamine more than that of either norepinephrine or 5-hydroxytryptamine (serotonin), even though histidine decarboxylase is less sensitive to the drug, simply because histidine decarboxylase is the only step and, hence, is rate-limiting in the biosynthesis of histamine. Dopa decarboxylase is not rate-limiting in the synthesis of either norepinephrine or 5-hydroxytryptamine until the enzyme is nearly completely inhibited. Another example of the determination of selectivity by the critical balance of the affected function is that of the mercurial diuretic drugs. An inhibition of only 1% in the tubular resorption of glomerular filtrate usually will double urine flow, since 99% of the glomerular filtrate is normally resorbed. Aside from the question of the possible concentration of diuretics in the urine, a drug-induced reduction of 1% in sulfhydryl enzyme activity in tissues other than the kidney usually is not accompanied by an observable change in function. Selectivity also can be determined by the pattern of distribution of destructive or activating enzymes among the tissues and by other factors.

Dose—Even the uninitiated person knows that the *dose* of a drug is the amount administered. However, the appropriate dose of a drug is not some unvarying quantity, a fact sometimes overlooked by pharmacists, official committees and physicians. The practice of pharmacy is entrapped in a system of fixed-dose formulations, so that fine adjustments in dosage are often difficult to achieve. Fortunately, there is usually a rather wide latitude allowable in dosages. It is obvious that the size of the recipient individual should have a bearing upon the dose, and the physician may elect to administer the drug on a body-weight or surface-area basis rather than as a fixed dose. Usually, however, a fixed dose is given to all adults, unless the adult is exceptionally large or small. The dose for infants and children often is determined by one of several formulas which take into account age or weight, depending on the age group of the child and the type of action exerted by the drug. Infants, relatively, are more sensitive to many drugs, often because enzyme systems which destroy the drugs may not be developed fully in the infant.

The nutritional condition of the patient, the mental outlook, the presence of pain or discomfort, the severity of the condition being treated, the presence of secondary disease or pathology, genetic and many other factors affect the dose of a drug necessary to achieve a given therapeutic response or to cause an untoward effect (Chapter 67). Even two apparently well-matched normal persons may require widely different doses for the same intensity of effect. Furthermore, a drug is not always employed for the same effect and, hence, not in the same dose. For example, the dose of a progestin necessary for an oral contraceptive effect is considerably different from that necessary to prevent spontaneous abortion, and a dose of an estrogen for the treatment of the menopause is much too small for the treatment of prostatic carcinoma.

From the above it is evident that the wise physician knows that *the dose of a drug is "enough"* (ie, no rigid quantity but rather that which is necessary and can be tolerated) and individualizes the regimen accordingly. The wise pharmacist also will appreciate this dictum and recognize that official or manufacturer's recommended doses are sometimes quite narrowly defined and may be very wide of the mark. They should serve only as a useful guide rather than as an imperative.

Potency and Efficacy—The *potency* of a drug is the reciprocal of dose. Thus, it will have the units of persons/unit weight of drug or body weight/unit weight of drug, etc. Potency generally has little utility other than to provide a means of comparing the relative activities of drugs in a series, in which case *relative potency*, relative to some prototype member of the series, is a parameter commonly used among pharmacologists and in the pharmaceutical industry.

Whether a given drug is more potent than another has little bearing on its clinical usefulness, provided that the potency is not so low that the size of the dose is physically unmanageable or the cost of treatment is higher than with an equivalent drug. If a drug is less potent but more selective, it is the one to be preferred. Promotional arguments in favor of a more potent drug thus are irrelevant to the impor-

tant considerations that should govern the choice of a drug. However, it sometimes occurs that drugs of the same class differ in the maximum intensity of effect; that is, some drugs of the class may be less efficacious than others, irrespective of how large a dose is used.

Efficacy connotes the property of a drug to achieve the desired response, and *maximum efficacy* denotes the maximum achievable effect. Even huge doses of codeine often cannot achieve the relief from severe pain that relatively small doses of morphine can; thus, codeine is said to have a lower maximum efficacy than morphine. Efficacy is one of the primary determinants of the choice of a drug.

Dose–Effect Relationships

The importance of knowing how changes in the intensity of response to a drug vary with the dose is virtually self-evident. Both the physician, who prescribes or administers a drug, and the manufacturer, who must package the drug in appropriate dose sizes, must translate such knowledge into everyday practice. Theoretical or molecular pharmacologists also study such relationships in inquiries into mechanism of action and receptor theory (see page 702). It is necessary to define two types of relationships: (1) dose–intensity relationship—ie, the manner in which the intensity of effect in the individual recipient relates to dose—and (2) dose–frequency relationship—ie, the manner in which the number of responders among a population of recipients relates to dose.

Dose–Intensity of Effect Relationships—Whether the intensity of effect is determined *in vivo* (eg, the blood-pressure response to epinephrine in the human patient) or *in vitro* (eg, the response of the isolated guinea pig ileum to histamine), the dose–intensity of effect (often called dose-effect) curve usually has a characteristic shape, namely a curve that closely resembles one quadrant of a rectangular hyperbola.

In the dose–intensity curve depicted in Fig 35-2, the curve appears to intercept the x axis at 0 only because the lower doses are quite small on the scale of the abscissa, the smallest dose being 1.5×10^{-3} μg. Actually, the x intercept has a positive value, since a finite dose of drug is required to bring about a response, this lowest effective dose being known as the *threshold dose*. Statistics and chemical kinetics predict

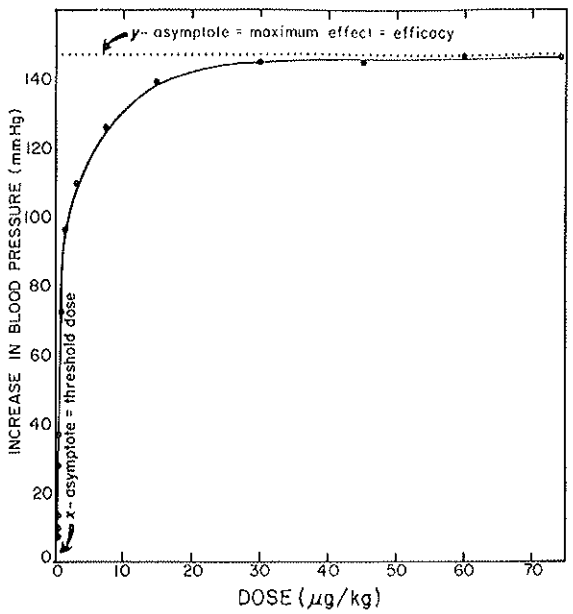


Fig 35-2. The relationship of the intensity of the blood-pressure response of the cat to the intravenous dose of norepinephrine.

that the curve should approach the y axis asymptotically. However, if the intensity of the measured variable does not start from zero, the curve possibly may have a positive y intercept (or negative x intercept), especially if the ongoing basal activity before the drug is given is closely related to that induced by the drug.

In practice, instead of an asymptote to the y axis, dose–intensity curves nearly always show an upward concave foot at the origin of the curve, so that the curve has a lopsided sigmoid shape. At high doses, the curve approaches an asymptote which is parallel to the x axis, and the value of the asymptote establishes the maximum possible response to the drug, or *maximum efficacy*. However, experimental data in the regions of the asymptotes generally are too erratic to permit an exact definition of the curve at the very low and very high doses. The example shown represents an unusually good set of data.

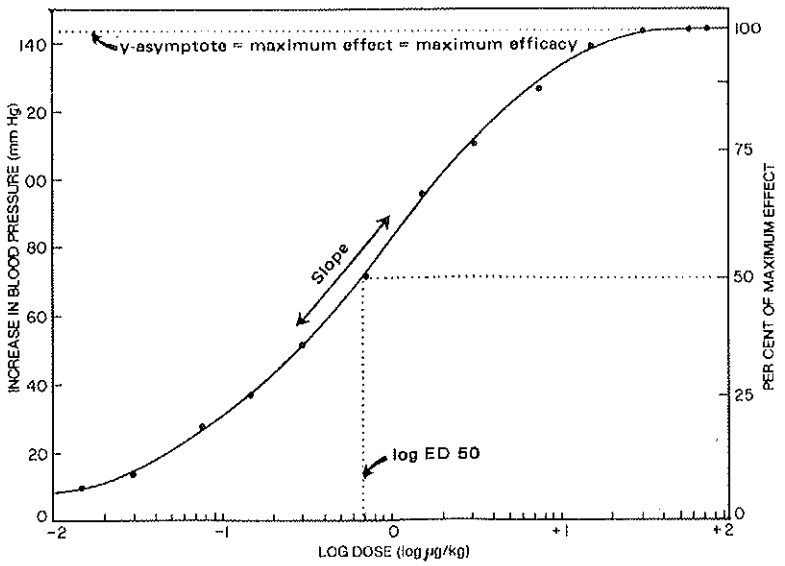


Fig 35-3. The relationship of the intensity of the blood-pressure response of the cat to the log of the intravenous dose of norepinephrine.

Because the dose range may be 100- or 1000-fold from the lowest to the highest dose, it has become the practice to plot dose–intensity curves on a logarithmic scale of abscissa; ie, to plot the log of dose versus the intensity of effect. Figure 35-3 is such a semilogarithmic plot of the same data as in Fig 35-2. In the figure the intensity of effect is plotted both in absolute units (at the left) or in relative units, as percent (at the right).

Although no new information is created by a semilogarithmic representation, the curve is stretched in such a way as to facilitate the inspection of the data; the comparison of results from multiple observations and the testing of different drugs also is rendered easier. In the example shown, the curve is essentially what is called a *sigmoid curve* and is nearly symmetrical about the point which represents an intensity equal to 50% of the maximal effect, ie, about the midpoint. The symmetry follows from the rectangular hyperbolic character of the previous Cartesian plot (Fig 35-2). The semilogarithmic plot reveals better the dose–effect relationships in the low-dose range, which are lost in the steep slope of the Cartesian plot. Furthermore, the data about the midpoint are almost a straight line; the nearly linear portion covers approximately 50% of the curve. The slope of the linear portion of the curve or, more correctly, the slope at the point of inflection, has theoretical significance (see *Drug Receptors and Receptor Theory*, page 702).

The upper portion of the curve approaches an asymptote, which is the same as that in the Cartesian plot. If the response system is completely at rest before the drug is administered, the lower portion of the curve should be asymptotic to the x axis. Both asymptotes and the symmetry derive from the law of mass action (see page 703).

Dose–intensity curves often deviate from the ideal configuration illustrated and discussed above. Usually, the deviate curve remains sigmoid but not extended symmetrically about the midpoint of the *linear* segment. Occasionally, other shapes occur, sometimes quite bizarre ones. Deviations may derive from multiple actions that converge upon the same final effector system, from varying degrees of metabolic alteration of the drug at different doses, from modulation of the response by feedback systems, from nonlinearity in the relationship between action and effect or from other causes.

It is frequently necessary to identify the dose which elicits a given intensity of effect. The intensity of effect that is generally designated is the 50% of maximum intensity. The corresponding dose is called the *50% effective dose*, or *individual ED50* (see Fig 35-3). The use of the adjective, *individual*, distinguishes the ED50 based upon the intensity of effect from the median effective dose, also abbreviated ED50, determined from frequency of response data in a population (see *Dose–Frequency Relationships*, this page).

Drugs that elicit the same quality of effect may be compared graphically. In Fig 35-4, five hypothetical drugs are compared. Drugs A, B, C and E all can achieve the same maximum effect, which suggests that the same effector system may be common to all. D possibly may be working through the same effector system, but there are no *a priori* reasons to think this is so. Only A and B have parallel curves and common slopes. Common slopes are consistent with, but in no way prove, the idea that A and B not only act through the same effector system but also by the same mechanism. Although drug–receptor theory (see *Drug Receptors and Receptor Theory*, page 702) requires that the curves of identical mechanism have equal slopes, examples of exceptions are known. Furthermore, mass-law statistics require that all simple drug-receptor interactions generate the same slope; only when slopes depart from this universal slope in accordance with distinctive characteristics of the response system do they provide evidence of specific mechanisms.

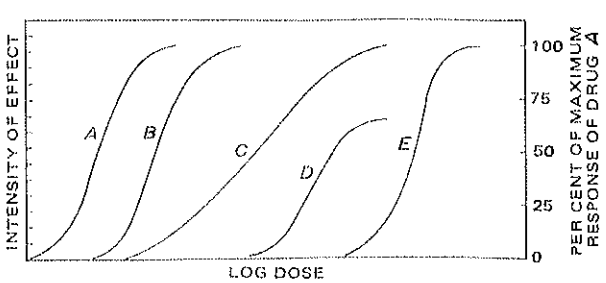


Fig 35-4. Log dose–intensity of effect curves of five different hypothetical drugs (see text for explanation).

The relative potency of any drug may be obtained by dividing the ED50 of the standard, or prototype, drug by that of the drug in question. Any level of effect other than 50% may be used, but it should be recognized that when the slopes are not parallel the relative potency depends upon the intensity of effect chosen. Thus, the potency of A relative to C (in Fig 35-4) calculated from the ED50 will be smaller than that calculated from the ED25.

The low maximum intensity inducible by D poses even more complications in the determination of relative potency than do the unequal slopes of the other drugs. If its dose–intensity curve is plotted in terms of percent of its own maximum effect, its relative inefficacy is obscured and the limitations of relative potency at the ED50 level will not be evident. This dilemma simply underscores the fact that drugs can be compared better from their entire dose–intensity curves than from a single derived number like ED50 or relative potency.

Drugs that elicit multiple effects will generate a dose–intensity curve for each effect. Even though the various effects may be qualitatively different, the several curves may be plotted together on a common scale of abscissa, and the intensity may be expressed in terms of percent of maximum effect; thus, all curves can share a common scale of ordinates in addition to common abscissa. Separate scales of ordinates could be employed, but this would make it harder to compare data.

The selectivity of a drug can be determined by noting what percent of maximum of one effect can be achieved before a second effect occurs. As with relative potency, selectivity may be expressed in terms of the ratio between the ED50 for one effect to that for another effect, or a ratio at some other intensity of effect. Similarly to relative potency, difficulties follow from nonparallelism. In such instances, selectivity expressed in dose ratios varies from one intensity level to another.

When the dose–intensity curves for a number of subjects are compared, it is found that they vary considerably from individual to individual in many respects; eg, threshold dose, midpoint, maximum intensity and sometimes even slope. By averaging the intensities of the effect at each dose, an average dose–intensity curve can be constructed.

Average dose–intensity curves enjoy a limited application in comparing drugs. A single line expressing an average response has little value in predicting individual responses unless it is accompanied by some expression of the range of the effect at the various doses. This may be done by indicating the standard error of the response at each dose. Occasionally, a simple scatter diagram is plotted in lieu of an average curve and statistical parameters (see Fig 10-21). An average dose–intensity curve also may be constructed from a population in which different individuals receive different doses; if sufficiently large populations are employed, the average curves determined by the two methods will approximate each other.

It is obvious that the determination of such average curves from a population sufficiently large to be statistically meaningful requires a great deal of work. Retrospective clinical data occasionally are treated in this way, but prospective studies infrequently are designed in advance to yield average curves. The usual practice in comparing drugs is to employ a quantal (all-or-none) end-point and plot the frequency or cumulative frequency of response over the dose range, as discussed below.

Dose–Frequency of Response Relationships—When an end-point is truly all-or-none, such as death, it is an easy matter to plot the number of responding individuals (eg, dead subjects) at each dose of drug or intoxicant. Many other responses that vary in intensity can be treated as all-or-none if simply the presence or absence of a response (eg, cough or no cough, convulsion or no convulsion) is recorded, without regard to the intensity of the response when it occurs.

When the response changes from the basal or control state in a less abrupt manner (eg, tachycardia, miosis, rate of gastric secretion) it may be necessary to designate arbitrarily some particular intensity of effect as the end-point. If the end-point is taken as an increase in heart rate of 20 beats/min, all individuals whose tachycardia is less than 20/min would be recorded as nonresponders, while all those with 20 or above would be recorded as responders. When the percent of responders in the population is plotted against the dose, a characteristic dose–response curve, more properly called a *dose–cumulative frequency* or *dose–percent* curve, is generated. Such a curve is, in fact, a cumulative frequency–distribution curve, the percent of responders at a given dose being the frequency of response.

Dose–cumulative frequency curves are generally of the same geometric shape as dose–intensity curves (namely, sigmoid) when frequency is plotted against log dose (see Fig 35-5). The tendency of the cumulated frequency of response (ie, percent) to be linearly proportional to the log of the dose in the middle of the dose range is called the *Weber–Fechner law*, although it is not invariable, as a true natural law should be. In many instances, the cumulative frequency is simply proportional to dose rather than log dose. The Weber–Fechner law applies to either dose–intensity or dose–cumulative frequency data. The similarity between dose–frequency and dose–intensity curves may be more than fortuitous, since the intensity of response will usually have an approximately linear relationship to the percent of responding *units* (smooth muscle cells, nerve fibers, etc) and, hence, is also a type of cumulative frequency of response. These are the same kind of statistics that govern the law of mass action.

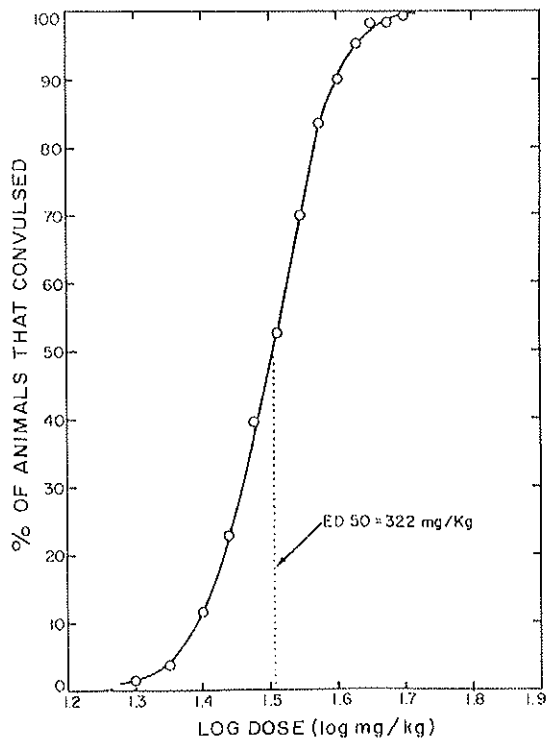


Fig 35-5. The relationship of the number of responders in a population of mice to the dose of pentylenetetrazol (courtesy, Drs DG McQuarry and EG Fingl, University of Utah).

If only the increase in the number of responders with each new dose is plotted, instead of the cumulative percent of responders, a bell-shaped curve is obtained. This curve is the first derivative of the dose–cumulative frequency curve and is a *frequency–distribution* curve (see Chapter 10). The distribution will be symmetrical—ie, *normal* or Gaussian (see Fig 10-5)—only if the dose–cumulative frequency curve is symmetrically hyperbolic. Because most dose–cumulative frequency curves are more nearly symmetrical when plotted semilogarithmically (ie, as log dose), dose–cumulative frequency curves are usually *log-normal*.

Since the dose–intensity and dose–cumulative frequency curves are basically similar in shape, it follows that the curves have similar defining characteristics, such as ED50, maximum effect (maximum efficacy) and slope. In dose–cumulative frequency data, the ED50 (*median effective dose*) is the dose to which 50% of the population responds (see Fig 35-5). If the frequency distribution is normal, the ED50 is both the arithmetic mean and median dose and is represented by the midpoint on the curve; if the distribution is log-normal, the ED50 is the median dose but not the arithmetic mean dose. The efficacy is the cumulative frequency summed over all doses; it is usually, but not always, 100%. The slope is characteristic of both the drug and test population. Even two drugs of identical mechanism may give rise to different slopes in dose–percent curves, whereas in dose–intensity curves the slopes are the same.

Statistical parameters (such as standard deviation), in addition to ED50, maximum cumulative frequency (efficacy) and slope, characterize dose-cumulative frequency relationships (see Chapter 10).

There are several formulations for dose–cumulative frequency curves, some of which are employed only to define the linear segment of a curve and to determine the statistical parameters of this segment. For the statistical treatment of dose–frequency data, see Chapter 10. One simple mathematical expression of the entire log-symmetrical sigmoid curve is

$$\log \text{dose} = K + f \log\left(\frac{\%\ \text{response}}{100\% - \text{response}}\right) \qquad (1)$$

where percent response may be either the percent of maximum intensity or the percent of a population responding. The equation is thus basically the same for both log normal dose–intensity and log normal dose–percent relationships. K is a constant that is characteristic of the midpoint of the curve, or ED50, and $1/f$ is characteristically related to the slope of the linear segment, which, in turn is closely related to the standard deviation of the derivative log-normal frequency-distribution curve.

The comparison of dose–percent relationships among drugs is subject to the pitfalls indicated for dose–intensity comparisons (see page 699), namely, that when the slopes of the curves are not the same (ie, the dose–percent curves are not parallel), it is necessary to state at which level of response a potency ratio is calculated. As with dose–intensity

data, potencies generally are calculated from the ED50, but potency ratios may be calculated for any arbitrary percent response. The expression of selectivity is, likewise, subject to similar qualifications, inasmuch as the dose–percent curves for the several effects are usually nonparallel.

The term *therapeutic index* is used to designate a quantitative statement of the selectivity of a drug when a therapeutic and an untoward effect are being compared. If the untoward effect is designated as *T* (for toxic) and the therapeutic effect as *E*, the therapeutic index may be defined as TD50/ED50 or a similar ratio at some other arbitrary levels of response. The TD and the ED are not required to express the same percent of response; some clinicians use the ratio TD1/ED99 or TD5/ED95, based on the rationale that if the untoward effect is serious, it is important to use a most-severe therapeutic index in passing judgment upon the drug. Unfortunately, therapeutic indices are known in man for only a few drugs.

There will be a different therapeutic index for each untoward effect that a drug may elicit, and, if there is more than one therapeutic effect, a family of therapeutic indices for each therapeutic effect. However, in clinical practice, it is customary to distinguish among the various toxicities by indicating the percent incidence of a given side effect.

Variations in Response and Responsiveness—From the above discussion of dose–frequency relationships and Chapter 10, it is obvious that in a normal population of persons there may be quite a large difference in the dose required to elicit a given response in the least-responsive member of the population and that to elicit the response in the most-responsive member. The difference ordinarily will be a function of the slope of the dose–percent curve, or, in statistical terms, of the standard deviation. If the standard deviation is large, the extremes of responsiveness of responders are likewise large.

In a normal population 95.46% of the population responds to doses within two standard deviations from the ED50 and 99.73% within three standard deviations. In log-normal populations the same distribution applies when standard deviation is expressed as log dose.

In the population represented in Fig 35-5, 2.25% of the population (two standard deviations from the median) would require a dose more than 1.4 times the ED50; an equally small percent would respond to 0.7 the ED50. The physician who is unfamiliar with statistics is apt to consider the 2.25% at either extreme as abnormal reactors. The statistician will argue that these 4.5% are within the normal population and only those who respond well outside of the normal population, at least three standard deviations from the median, deserve to be called abnormal.

Irrespective of whether the criteria of abnormality that the physician or the statistician obtain, the term *hyporeactive* applies to those individuals who require abnormally high doses and *hyperreactive* to those who require abnormally low doses. The terms *hyporesponsive* and *hyperresponsive* also may be used. It is incorrect to use the terms hyposensitive and hypersensitive in this context; *hypersensitivity* denotes an allergic response to a drug and should not be used to refer to hyperreactivity. The term *supersensitivity* correctly applies to hyperreactivity that results from denervation of the effector organ; it is often more definitively called denervation supersensitivity. Sometimes hyporeactivity is the result of an immunochemical deactivation of the drug, or *immunity*. Hyporeactivity should be distinguished from an increased dose requirement that results from a severe pathological condition. Severe pain requires large doses of analgesics, but the patient is not a hyporeactor; what has changed is the baseline from which the endpoint quantum is measured. The responsiveness of a patient to certain drugs sometimes may be determined by the history of previous exposure to appropriate drugs.

Tolerance is a diminution in responsiveness as use of the drug continues. The consequence of tolerance is an increase in the dose requirement. It may be due to an increase in the rate of elimination of drug (as discussed elsewhere in this chapter), to reflex or other compensatory homeostatic adjustments, to a decrease in the number of receptors or in the number of enzyme molecules or other coupling proteins in the effector sequence, to exhaustion of the effector system or depletion of mediators, to the development of immunity or to other mechanisms. Tolerance may be gradual, requiring many doses and days to months to develop, or acute, requiring only the first or a few doses and only minutes to hours to develop. Acute tolerance is called *tachyphylaxis*.

Drug resistance is the decrease in responsiveness of microorganisms, neoplasms or pests to chemotherapeutic agents, antineoplastics or pesticides, respectively. It is not tolerance in the sense that the sensitivity of the individual microorganism or cancer cell decreases; rather, it is the survival of normally unresponsive cells which then pass the genetic factors of resistance on to their progeny.

Patients who fail to respond to a drug are called *refractory*. Refractoriness may result from tolerance or resistance, but it also may result from the progression of pathological states that negate the response or render the response incapable of surmounting an overwhelming pathology. Rarely, it may result from a poorly developed receptor or response system.

Sometimes a drug evokes an unusual response that is *qualitatively* different from the expected response. Such an unexpected response is called a *meta-reaction*. A not uncommon meta-reaction is a central nervous stimulant rather than depressant effect of phenobarbital, especially in women. Pain and certain pathological states sometimes favor *meta*-reactivity. Responses that are different in infants or the aged than in young and middle-aged people are not *meta*-reactions if the response is usual in the age group. The term *idiosyncrasy* also denotes *meta*-reactivity, but the word bas been so abused that it is recommended that it be dropped. Although hypersensitivity may cause unusual effects, it is not included in *meta*-reactivity.

Drug Receptors and Receptor Theory

Most drugs act by combining with some key substance in the biological milieu that has an important regulatory function in the target organ or tissue. This biological partner of the drug goes by the name of *receptive substance* or *drug receptor*. The receptive substance is considered mostly to be a cellular constituent, although in a few instances it may be extracellular, as the cholinesterases are, in part. The receptive substance is thought of as having a special chemical affinity and structural requirements for the drug. Drugs such as emollients, which have a physical rather than chemical basis for their action, obviously do not act upon receptors. Drugs such as demulcents and astringents, which act in a nonselective or nonspecific chemical way, also are not considered to act upon receptors, since the candidate receptors have neither sharp chemical nor biological definition. Even antacids, which react with the extremely well-defined hydronium ion, cannot be said to have a receptor, since the reactive proton has no permanent biological residence.

Because of early preoccupation with physical theories of action and the classical and illogical dichotomy of chemical

and physical molecular interaction, there is a reluctance to admit receptors for drugs such as local anesthetics, general anesthetics, certain electrolytes, etc, which generally are not accepted to combine selectively with distinct cellular or organelle membrane constituents. The word receptor often is used inconsistently and intuitively. However, the term is a legitimate symbol for that biological structure with which a drug interacts to initiate a response. Ignorance of the identities of many receptors does not detract from, but rather increases, the importance of the term and general concept.

Once a receptor is identified, it frequently is no longer thought of as a receptor, although such identification may afford the basis of profound advances in receptor theory. Since the effects of anticholinesterases are derived only indirectly from inhibition of cholinesterase and no drugs are known that stimulate the enzyme, it may be argued that it is not a receptor. Nevertheless, a number of drugs ultimately act indirectly through the inhibition of such modulator enzymes and it is important for the theoretician to develop models based upon such indirect interrelations.

Enzymes, of course, readily suggest themselves as candidates for receptors. However, there is more to cellular function than enzymes. Receptors may be membrane or intracellular constituents that govern: the spatial orientation of enzymes, gene expression, compartmentalization of the cytoplasm, contractile or compliant properties of subcellular structures or permeability and electrical properties of membranes. For nearly every cellular constituent there can be imagined a possible way for a drug to affect its function; therefore, few cellular constituents can be dismissed *a priori* as possible receptors. All the receptors for neurotransmitters and autonomic agonists are membrane proteins with agonist-binding groups projecting into the extracellular space. The transducing apparatus, whereby an occupied receptor elicits a response, is called a *coupling system*. Excitatory neurotransmitters in the central nervous system, and nicotinic receptors elsewhere, are coupled to ion channels which, when opened, permit the rapid ingress, especially of sodium ions. GABA (γ-amino-butyric acid) and glycine are coupled to inhibitory chloride channels. Benzodiazepine receptors are coupled to the GABA-receptor. Beta-adrenergic receptors and a number of receptors for polypeptide hormones interact with a stimulatory GDP/GTP-binding protein (G-protein) which can activate the enzyme adenylate cyclase. The cyclase then produces 3′,5′-cyclic AMP (cAMP) which, in turn, activates protein kinases. Other receptors interact with inhibitory G-proteins. Some receptors couple to guanylate cyclase.

Alpha-adrenergic, some muscarinic and various other receptors couple to the membrane enzyme, phospholipase-C, which cleaves inositol phosphates from phosphoinositides. The cleavage product, 1,4,5-inositol triphosphate (IP_3), then causes an increase in intracellular calcium, whereas the product, diacylglycerol (DAG), activates kinase-C. There are a number of other less ubiquitous coupling systems. Substances such as cAMP, cGMP, IP_3 and DAG are called *second messengers*.

It has been found that there may be several different receptors for a given agonist. Differences may be shown not only in the types of coupling systems and effects but also by differential binding of agonists and antagonists, desensitization kinetics, physical and chemical properties, genes and amino acid sequences. The differentiation among receptor subtypes is called *receptor classification*. Receptor subtypes are designated by Greek or Arabic alphabetical prefixes and/or numerical subscripts. There are at least two each of beta-adrenergic, histaminergic, serotoninergic, GABAergic and benzodiazepine receptors, probably three of muscarinic and alpha-adrenergic and five of opioid receptor subtypes.

Occupation and Other Theories

Drug-receptor interactions are governed by the law of mass action, a concept initiated by Langley in 1878. However, most chemical applications of mass law are concerned with the rate at which reagents disappear or products are formed, whereas receptor theory usually concerns itself with the fraction of the receptors combined with a drug, similar to theories of adsorption. The usual concept is that only when the receptor actually is occupied by the drug is its function transformed in such a way as to elicit a response. This concept has become known as the *occupation theory*. The earliest clear statement of its assumptions and formulations is often credited to Clark in 1926, but both Langley and Hill made important contributions to the theory in the first two decades of this century.

In all receptor theories, the terms agonist, partial agonist and antagonist are employed. An *agonist* is a drug that combines with a receptor to initiate a response.

In the classical occupation theory, two attributes of the drug are required: (1) *affinity*, a measure of the equilibrium constant of the drug-receptor interaction, and (2) *intrinsic activity*, or *intrinsic efficacy* (not to be confused with efficacy as intensity of effect), a measure of the ability of the drug to induce a positive change in the function of the receptor.

A *partial agonist* is a drug that can elicit some but not a maximal effect and which antagonizes an agonist. In the occupation theory it would be a drug with a favorable affinity but a low intrinsic activity.

A *competitive antagonist* is a drug that occupies a significant proportion of the receptors and thereby preempts them from reacting maximally with an agonist. In the occupation theory the prerequisite property is affinity without intrinsic activity.

A *noncompetitive antagonist* may react with the receptor in such a way as not to prevent agonist–receptor combination but to prevent the combination from initiating a response, or it may act to inhibit some subsequent event in the chain of action–effect–action–effect that leads to the final overt response.

The mathematical formulation of the receptor theories derives directly from the law of mass action and chemical kinetics. Certain assumptions are required to simplify calculations. The key assumption is that the intensity of effect is a direct linear function of the proportion of receptors occupied. The correctness of this assumption is most improbable on the basis of theoretical considerations, but empirically it appears to be a close enough approximation to be useful. A second assumption upon which formulations are based is that the drug-receptor interaction is at equilibrium. Another common assumption is that the number of molecules of receptor is negligibly small compared to that of the drug. This assumption is undoubtedly true in most instances, and departures from this situation greatly complicate the mathematical expression of drug-receptor interactions.

The first clearly stated mathematical formulation of drug–receptor kinetics was that of Clark.[1] In his equation,

$$Kx^n = \frac{y}{100 - y} \tag{2}$$

where K is the affinity constant, x is the concentration of drug, n is the molecularity of the reaction, and y is the percent of maximum response. Clark assumed that y was a linear function of the percent of receptors occupied by the drug, so that y could also symbolize the percent of receptors occupied. When the equation is rearranged to solve for y,

$$y = \frac{100Kx^n}{1 + Kx^n} \tag{3}$$

A Cartesian plot of this equation is identical in form to that shown in Fig 35-2. When y is plotted against log x instead of x, the usual sigmoid curve is obtained. Thus, it may be seen that the dose–intensity curve derives from mass action equilibrium kinetics, which in turn derive from the statistical nature of molecular interaction. The fact that dose–intensity and dose–percent curves have the same shape shows that they involve similar statistics.

If Eq 2 is put into log form

$$\log K + n \log x = \log \frac{y}{100 - y} \qquad (4)$$

a plot of log $y/100 - y$ against log x then will yield a straight line with a slope of n; n is theoretically the number of molecules of drug which react with each molecule of receptor. At present, there are no known examples in which more than one molecule of agonist combines with a single receptor, hence, n should be equal to 1, universally. Nevertheless, n often deviates from 1; deviations occur because of cooperative interactions among receptors (*cooperativity*), *spare receptors* (see below), amplifications in the response system (*cascades*), receptor coupling to more than one sequence (eg, to both adenylate cyclase and calcium channels) and other reasons. In these departures from $n = 1$, the slope becomes a characteristic of the mechanism of action and response system.

The probability that a molecule of drug will react with a receptor is a function of the concentration of both drug and receptor. The concentration of receptor molecules cannot be manipulated like the concentration of a drug. But, as each molecule of drug combines with a receptor, the population of free receptors is diminished accordingly. If the drug is a competitive antagonist, it will diminish the probability of an agonist–receptor combination in direct proportion to the percent of receptor molecules preempted by the antagonist. Consequently, the intensity of effect will be diminished. However, the probability of agonist–receptor interaction can be increased by increasing the concentration of agonist, and the intensity of effect can be restored by appropriately larger doses of agonist. Addition of more antagonist will again diminish the response, which can, again, be overcome or *surmounted* by more agonist.

Clark showed empirically, and by theory, that as long as the ratio of antagonist to agonist was constant, the concentration of the competitive drugs could be varied over an enormous range without changing the magnitude of the response (see Fig 35-6). Since the presence of competitive antagonist only diminishes the probability of agonist–receptor combination at a given concentration of agonist and does not alter the molecularity of the reaction, it also follows that the effect of the competitive antagonist is to shift the dose–intensity curve to the right in proportion to the amount of antagonist present; neither shape nor slope of the curve is changed (see Fig 35-7). Both Figs 35-6 and 35-7 are from Clark's original paper on competitive antagonism.[1]

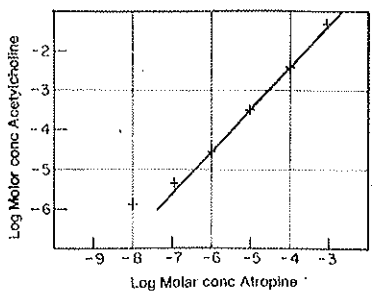


Fig 35-6. Direct proportionality of the dose of agonist (acetylcholine) to the dose of antagonist (atropine) necessary to cause a constant degree of inhibition (50%) of the response of the frog heart (courtesy, adaptation, Clark[1]).

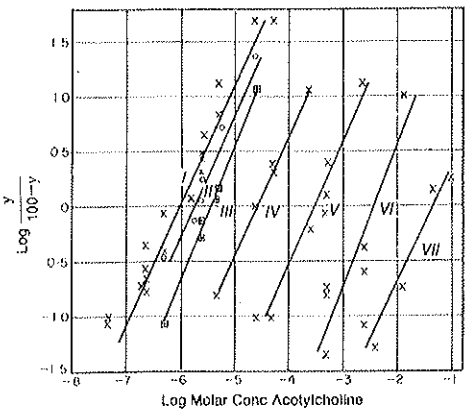


Fig 35-7. Effect of an antagonist to shift the log dose–intensity curve to the right without altering the slope. The effector is the isolated heart. *I:* no atropine; *II:* atropine, $10^{-8}M$; *III:* $10^{-7}M$; *IV:* $10^{-6}M$; *V:* $10^{-5}M$; *VI:* $10^{-4}M$; *VII:* $10^{-3}M$. *Y:* % of maximum intensity of response; the function log $y/(100\text{-}y)$ converts the log dose–intensity relationship to a straight line (courtesy, adaptation, Clark[1]).

Many refinements of the Clark formula have been made, but they will not be treated here; details and citations of relevant literature can be found among various works on receptors cited in the Bibliography. Several refinements are introduced to facilitate studies of competitive inhibition. The introduction of the concepts of intrinsic activity[2] and efficacy[3] required appropriate changes in mathematical treatment.

Another important concept has been added to the occupation theory, namely the concept of *spare receptors*. Clark assumed the maximal response to occur only when the receptors were completely occupied, which does not account for the possibility that the maximum response might be limited by some step in the action–effect sequence subsequent to receptor occupation. Work with isotopically labeled agonists and antagonists and with dose–effect kinetics has shown that the maximal effect sometimes is achieved when only a small fraction of the receptors are yet occupied. The mathematical treatment of this phenomenon has enabled theorists to explain several puzzling observations that previously appeared to contradict occupation theory.

The classical occupation theory fails to explain several phenomena satisfactorily, and it is unable to generate a realistic model of intrinsic activity and partial agonism. A rate theory, in which the intensity of response is proportional to the rate of drug–receptor interaction instead of occupation, was proposed to explain some of the phenomena that occupation theory could not, but the rate theory was unable to provide a realistic mechanistic model of response generation, and it had other serious limitations as well.

The phenomena that neither the classical occupation nor rate theory could explain can be explained by various theories in which the receptor can exist in at least two conformational states, one of which is the active one; the drug can react with one or more conformers. In a *two-state* model[4]

$$R \rightleftharpoons R^*$$

where R is the inactive and R* is the active conformer. The agonist combines mainly with R*, the partial agonist can combine with both R and R* and the antagonist can combine with R, the equilibrium being shifted according to the extent of occupation of R and R*. Other variations of occupation theory treat the receptor as an aggregate of subunits which interact cooperatively.[5]

The Nature of Receptor Groups and Models of Receptors

A *receptor group* is that portion of the receptor molecule with which an agonist acts and which is vital to the function of the receptor. Studies of receptor group composition and configuration are too complex for the purposes of this text; consequently, only a brief sketch will be made here to orient the reader to the nature of the approach.

From the chemical configuration and reactivity of agonists and antagonists, certain deductions can be made about the structure of a receptor group. For example, all highly active agonists of muscarinic receptors are cations at physiological pH. This suggests that the receptor group contains an anionic group and that the force of attraction is electrostatic, at least in part, which agrees with thermodynamic data. That van der Waals forces (especially Heitler-London fluctuation forces) may also make an important contribution to binding is suggested by the requirement for *N*-methyl groups and by the low but definite activity of the nonionizable quaternary carbon analog of acetylcholine, 3,3-dimethylbutyl acetate. This establishes a requirement for an auxiliary structure close to the anionic site. Studies of the contribution to activity of ester and carbonyl oxygen among analogs of acetylcholine, intramolecular distances and the stereospecificity of various isomers and conformers have indicated a partial cationic (proton donor) site between 2.5 and 4 Å and a region of high electronic density (electron donor) between 5 and 7 Å from the anionic site. This is similar to the way in which the active site of acetylcholinesterase was mapped (see page 427, and Figs 25-44, 45 and 46).

The structure–activity relationships among competitive inhibitors also must be consistent with any model of a receptor. However, binding sites additional to the receptor group can be involved, and results are frequently more difficult to interpret than those with agonists. Nevertheless, studies with antagonists have made a substantial contribution to receptor group analysis. There is considerable interest in antagonists that combine irreversibly with the receptor, since such drugs offer a way of marking (affinity labeling) the receptor for isolation and for identification of the receptor group.

Since receptors for autonomic agonists are embedded in the cell membrane, they have been difficult to isolate without inactivation. Several laboratories have succeeded in isolating proteins, the chemical properties of which are consistent with those expected of various receptors. Receptors for steroid hormones have been easier to isolate, and some have been characterized relatively well. Further details of drug-receptor interactions and the nature of receptors can be found in the works on receptors and molecular pharmacology.

Up- and Down-Regulation—In many receptor-effector systems, if there is a paucity of agonist, the system will respond by increasing the responsiveness, number of receptors on the effector membrane or number of coupling proteins or enzymes in the effector system. This is known as *up-regulation*. In adrenergic systems, sympathetic denervation has been shown to increase the number of post-synaptic β-adrenoreceptors at some junctions and the availability of nucleotide-binding protein units and/or adenylate cyclase molecules at others. Hyperthyroid activity also increases the number of β-adrenoreceptors in heart muscle, which explains the excessive heart rate. Denervation of skeletal muscle causes a great multiplication of what is normally a minor type of nicotinic receptor, and the new receptors spread across the entire myocyte membrane. Prolonged blockade of receptors by antagonists also may cause up-regulation. The abrupt discontinuation of treatment, such that drug levels fall faster than re-regulation, may be followed by excessive activity, eg, in pernicious tachycardia and angina pectoris from abrupt withdrawal of propranolol.

Excessive agonism will lead to a decrease in the number of receptors or in stimulus-response coupling. This is one cause of tachyphylaxis or tolerance, such as occurs to the bronchodilator effects of β-adrenoreceptor agonists. Abrupt withdrawal may result in poor residual function or in rebound effects, depending upon the type of effect caused by the agonist. Excessive agonism also may cause desensitization by agonist-induced changes in receptor conformation to inactive, slowly reconformable states.

Mechanism of Drug Action

Any metabolic or physiological function provides a potential mechanism of action of a drug. The term *mechanism of action* has been employed in a number of ways. In the past it was often the habit to confuse the site, or locus of action, with the mechanism of action. For example, the mechanism of the hypotensive action of tetraethylammonium ion originally was described as that of ganglionic blockade, which did nothing more than identify the anatomical structure upon which the drug acted. In a general sense, this was a partial elucidation of the mechanism of action, if mechanism is used in the mechanical sense of the entire linkage between the input and output of a machine. However, there has been a gradual narrowing of the definition of mechanism of action to be restricted to only the first event in the action–effect sequence, that is, only to the alteration of receptor function by the drug. In this sense, the mechanism of action of tetraethylammonium is defined more appropriately as that of competition with acetylcholine for nicotine cholinergic receptors on the postsynaptic ganglion cell membrane, even though the alteration in receptor function is not defined. The ultimate mechanism of action is known for only a few drugs.

It is customary to speak of a drug as a stimulant or a depressant, of the action as being excitatory or inhibitory, etc. Such terms describe only the effect and not the action, and they have no bearing upon whether the drug augments receptor function or diminishes it. In biological systems, positive and negative modulation and feedback occur at every level, the organ as well as the subcellular. Thus, an agonist to a negative modulator may be able to bring about the same effect as an antagonist to a positive modulator. It is possible for an antagonist or inhibitor to elicit an excitatory effect. An example is the convulsant action of strychnine, which results from its antagonism of glycine, an important mediator of postsynaptic inhibition in the central nervous system. Conversely, it is possible for an agonist to elicit an inhibitory effect. An example is the reflex bradycardia that results from the stimulant action of veratrum alkaloids on chemoreceptors in the left ventricle.

Because of the central role *enzymes play* in cellular function, it is not surprising that thoughts about the mechanism of action of drugs has focused largely upon enzymes. Agonist drugs conceivably could serve as substrates, cofactors or activators. At the present time, no drug is known definitely to exert its action as a substrate or as a cofactor, exclusive of vitamins and known nutrients. However, at least three classes of drugs are known and several are suspected to work through the activation of enzymes.

The most notable example of enzyme activation is that of epinephrine and similar β-adrenoreceptor agonists, which activate adenyl cyclase to increase the production of 3′,5′-cyclic adenylic acid (cyclic AMP; cAMP). The metabolic and cardiac effects of catecholamines are attributable, in part, to the increment in cAMP. One modulator of adenyl cyclase is the β-adrenergic receptor. The β-adrenoreceptor is coupled to adenylate cyclase through a regulatory protein that binds GDP and GTP (G-protein). When GDP is present, the agonist-receptor complex is associated with the regulatory protein. GTP causes transfer of the regulatory protein to adenylate cyclase and dissociation of the β-adrenoreceptor. Glucagon also owes its hyperglycemic action to activation of hepatic adenylate cyclase. A number of other agonists also activate adenylate cyclase. There is, thus, the interesting phenomenon of one enzyme, adenyl cyclase, being activated by numerous chemically unrelated drugs. Since β-adrenergic-blocking agents do not antagonize glucagon, it is obvious that glucagon works upon a different receptor than does epinephrine.

Thus, cAMP activates protein kinases that increase the activity of phosphorylase, actomyosin, the sequestration of calcium by the sarcoplasmic reticulum and calcium channels. Therefore, a brief activation of the β-adrenoreceptor sets in motion a cascade of events that greatly amplify the signal. Kinases also participate in down-regulation and desensitization.

Other important enzymes coupled to receptors are guanylate cyclase and phospholipases A and C, which are involved with membrane fluidity and calcium channels, respectively.

Many drugs are inhibitors of enzymes. When the drug is a *competitive inhibitor* of a natural endogenous substrate of the enzyme, it is called an *antimetabolite* (see also page 431). Examples of antimetabolites are sulfonamides, which compete with para-aminobenzoic acid and, thus, interfere with its incorporation into dihydrofolic acid and methotrexate, which competes with folic acid for dihydrofolate reductase and, thus, interferes with the formation of folinic acid. It might seem that anticholinesterases are also antimetabolites, although they are never placed into that classification. The reason is that the products of cholinesterase–acetylcholine interaction do not subserve important metabolic functions, as do folic and folinic acids, so that the organism is not deprived of an important metabolite by the action of the cholinesterase inhibitors.

Some drugs are competitive inhibitors of enzyme systems whose natural function appears not to produce useful metabolites but to rid the body of foreign substances. Inhibitors of the hepatic microsomes and probenecid fall into this category; the hepatic microsomes do perform a few biotransformations on endogenous substrates, but the renal tubular anion transport system does not appear to be required to eliminate any important endogenous substances.

Since neither the hepatic microsomes nor the tubular anion transport system seems to be involved in response systems, inhibitors of these enzyme systems are antagonists without corresponding agonists. Indeed, even natural endogenous substrates of enzymes are rarely considered to be agonists.

Noncompetitive enzyme inhibitors among drugs also are known. Examples are cyanide, fluoride, disulfiram and cardiac glycosides. When enzyme inhibition brings about a positive response—eg, the cholinergic effects of the anticholinesterases or the effects of diazoxide consequent to inhibition of phosphodiesterase—the drug appears to be an agonist. Yet, there can be no competitive antagonist to such an inhibitor, since the competitor to the drug is more substrate, to which the effect of the drug is actually attributable.

Acetylcholine increases the permeability of the subsynaptic membrane to cations and the heart muscle membrane to potassium. The mechanism is thought generally to involve a change in conformation of a protein constituent of the potassium channel, so that pore size or permeability constant is affected. The muscarinic receptor is coupled to the potassium channel through a G-protein. Other autonomic agonists also are known to alter the permeability to ions, in part through activation of adenyl cyclase, guanyl cyclase, phospholipase-c or other enzymes. Many drugs and toxins act through *alterations in the structural and physical properties of membranes*. To the extent that some of such substances may disperse themselves generally throughout the lipid phase of the membrane rather than to combine with special chemical entities, no definite receptors for such drugs can be said to exist.

The mechanism of action of certain drugs, especially autonomic drugs, often is stated to be *mimicry* of a natural neurohumor or hormone. Thus, methacholine mimics acetylcholine as an agonist. This does not define the mechanism of action, unless the mechanism of action of the natural substance is known.

Mimicry usually occurs because of a structural similarity between the natural substance and the mimetic drug. Mimicry in agonist functions is easy to demonstrate, but the site of action may not always be mimicry of the natural agonist at its receptor but rather at an allosteric site on a receptor or at its storage site to *release* the natural agonist.

Examples of mimetics that act by release of the natural mediator are indirectly acting sympathomimetics such as *d*-amphetamine, mephentermine, ephedrine (in part), tyramine and others, which are now known to act by displacing norepinephrine from storage sites within the adrenergic neuron. Many of such indirectly acting sympathomimetics lack a direct action on the adrenergic receptor, although some, like ephedrine, act both upon the receptor and the storage complex. Another mimetic by a release mechanism is carbachol, which promotes the presynaptic discharge of acetylcholine.

In these examples, there is a close structural similarity between the mimetic and the released mediator. In the case of many releasers of histamine (such as tubocurarine, polymyxin or morphine), no close chemical relationship exists between the releaser and the released. In such instances, release has been explained by activation of receptors on the mast-cell membrane which promote exocytosis of the histamine-containing granules, by an influx of calcium and activation of microtubules, all of which may be involved in moving the granules out of the mast cell.

Structural similarity also may aid mimicry by promoting chemical combination with an enzyme of destruction or some other means of disposition. For example, metaraminol, amphetamine, etc *inhibit membrane transport* into the neuron and, hence, inhibit the neuronal recapture of released norepinephrine. Consequently, the extraneuronal concentration of norepinephrine in the nearby region of the receptors does not drop as rapidly as in the absence of the mimetic, and the action of the mediator is sustained.

Some inhibitors of the enzymes of the destruction of mediators are structurally similar enough to the mediator to have some agonist action. This is true of neostigmine, which has a direct stimulant action on nicotinic receptors in addition to its anticholinesterase action. In contrast, the anticholinesterase, physostigmine, has some antagonist actions on cholinergic receptors and also an effect to interfere with acetylcholine synthesis.

The above multiple actions come about because all the structures that interact with a small molecule mediator (the receptor, synthesizing enzyme, destructive enzyme, storage molecule, membrane transport carrier) must have some common structural features and affinities. A drug that re-

acts with one of these molecules has a distinct probability of interacting with another.

The recognition of the critical role of *ions* in the function of membranes, the excitability of cells and the activity of many enzymes has generated a renewed interest in ions in the mechanism of action of certain drugs. The inorganic ions, some of which are used as drugs, lend themselves automatically to a discussion of ionic mechanisms. The repair of electrolyte deficiencies by replacement therapy warrants no further comment here. Some nonphysiological ions act as imperfect impersonators of physiological ions; lithium partly substitutes for sodium, bromide for chloride and thiocyanate for iodide, and each may owe its pharmacological action, in part, to a sluggish mobility through membrane channels, through which their sister ions normally pass readily when traffic is not impeded by "slowly moving vehicles." Iodide has an effect to increase the penetrance of drugs into caseous and necrotic areas, to aid in the resolution of gummatous lesions, to reduce the viscosity of mucous secretions and other odd effects; it is thought to do so by increasing the hydration of collagen and mucoproteins by a poorly understood mechanism. The transition elements and heavy metals have in common the ability to form complexes with a variety of physiologically active substances, particularly the active centers of many enzymes. *Chelation* and other types of *complexation* are the mechanisms of action of several drugs used to treat heavy-metal intoxication, diseases that involve abnormal body burdens or plasma levels of heavy metals and hypercalcemia. Chelates and chelation are discussed in more detail in Chapter 14.

There is much interest in the effects of drugs on ion movements. Cardiac glycosides are known to inhibit an ATPase involved in the membrane transport of sodium and several other substances, which indirectly causes an increase in intracellular calcium content. In part, the mechanisms of action of local anesthetics, quinidine and various other drugs also are speculated to involve calcium movements. In the past decade there has appeared a whole new class of drugs, the calcium channel blockers.

Concomitant with the development of molecular biology was the appreciation that drugs act through *nuclear* and *extranuclear genetic mechanisms.* Nitrogen mustards have long been known to interfere with the replication of DNA. Streptomycin, kanamycin, neomycin and gentamicin cause misreading by the ribosomes of the code incorporated into messenger RNA; tetracyclines, erythromycin and chloramphenicol inhibit the synthesis of protein at the ribosomes; and chloroquine, novobiocin and colchicine inhibit DNA polymerase. Other drugs induce the production of enzymes; aldosterone appears to act by inducing the synthesis of the enzyme, membrane ATPase, necessary to sodium transport. In general, steroid hormones combine with a cytosolic receptor, the complex of which is processed and translocated to the chromatin, where gene expression is altered. Many drugs induce one or more of the hepatic and extrahepatic cytochrome P-450 enzymes.

A number of drugs have simple mechanisms that do not involve an action at the cellular level. Examples are bulk and saline cathartics, osmotic diuretics and cholestyramine. Although such drugs usually do not generate much excitement among pharmacologists, they do serve as a reminder of the many avenues through which a mechanism of action may be expressed. Throughout the various chapters of Part 6, specific mechanisms of action may be mentioned.

Absorption, Distribution and Excretion

No matter by which route a drug is administered it must pass through several to many biological membranes during the processes of absorption, distribution, biotransformation and elimination. Since membranes are traversed in all of these events, the subject of this section will begin with a brief description of biological membranes and membrane processes and the relationship of the physiochemical properties of a drug molecule to penetrance and transport.

Structure and Properties of Membranes

The concept that a membrane surrounds each cell arose shortly after the cellular nature of tissue was discovered. The biological and physiochemical properties of cells seemed in accord with this view. In the past, from time to time, the actual existence of the membrane has been questioned by brilliant men, and ingenious explanations have been advanced to explain cellular integrity and the osmotic and electrophysiological properties of cells. Microchemical, x-ray diffraction, electron microscopic, nuclear magnetic resonance, electron spin resonance and other investigations have proved both the existence and nature of the plasma, mitochondrial, nuclear and other cell membranes. The description of the plasma membrane that follows is much oversimplified, but it will suffice to provide a background for an understanding of penetrance into and through membranes.

Structure and Composition—The cell membrane has been described as a "mayonnaise sandwich," in which a bimolecular layer of lipid material is entrained between two parallel monomolecular layers of protein. However, the protein does not make continuous layers, like the bread in a sandwich, but rather is sporadically scattered over the surfaces, like icebergs; ie, much of the protein is below the surface. In Fig 35-8 the lipid layers are represented as a somewhat orderly, closely packed lamellar array of phospholipid molecules associated tail-to-tail, each "tail" being an alkyl chain or steroid group and the "heads" being polar groups, including the glycerate moieties, with their polar ether and carbonyl oxygens and phosphate with attached polar groups. In reality, the lamellar portion is probably not

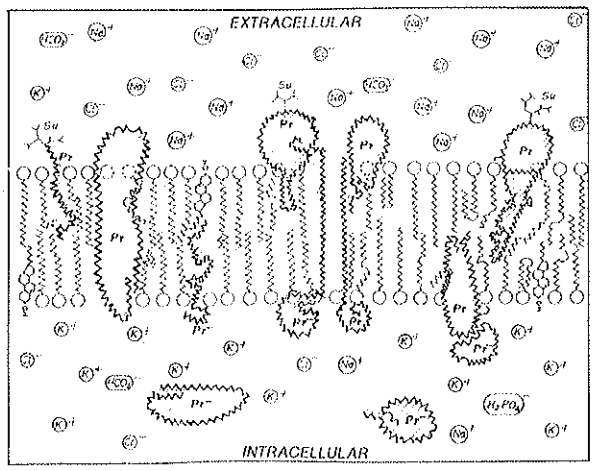


Fig 35-8. Simplified cross section of a cell membrane (components are not to scale). The lipid interior of the lamellar portion of the membrane consists of various phospholipids, fatty acids, cholesterol and other steroids. Ions are indicated in order to illustrate differences in size relative to the channel. *Pr:* protein; *Su:* sugar.

so orderly, since its composition is quite complex. Chains of fatty acids of different degrees of saturation and cholesterol cannot array themselves in simple parallel arrangements. Furthermore, the polar heads will assume a number of orientations depending upon the substances and groups involved. Moreover, the lamellar portion is penetrated by large globular proteins, the interior of which, like the lipid layers, has a high hydrophobicity, and some fibrous proteins.

The plasma membrane appears to be asymmetrical. The lipid composition varies from cell type to cell type and perhaps from site to site on the same membrane. There are, for example, differences between the membrane of the endoplasmic reticulum and the plasma membrane, even though the membranes are coextensive. Where membranes are double, the inner and outer layers may differ considerably; the inner and outer membranes of mitochondria have been shown to have strikingly different compositions and properties. Some authorities have expressed doubt as to the existence of the protein layers in biological membranes, although the evidence is preponderantly in favor of at least an outer glycoprotein coat. Sugar moieties also are attached to the outer proteins; these sugar moieties are important to cellular and immunological recognition and adhesion and have other functions as well.

The cell membrane appears to be perforated by water-filled pores of various sizes, varying from about 4 to 10 Å, the majority of which are about 7 Å. Probably all major ion channels are through the large globular proteins that traverse the membrane. Through these pores pass inorganic ions and small organic molecules. Since sodium ions are more hydrated than potassium and chloride ions, they are larger and do not pass as freely through the pores as potassium and chloride. The vascular endothelium appears to have pores at least as large as 40 Å, but these seem to be interstitial passages rather than transmembrane pores. Lipid molecules small enough to pass through the pores may do so, but they have a higher probability of entering into the lipid layer, from where they will equilibrate chemically with the interior of the cell. From work on monolayers, some researchers contend that it is not necessary to postulate pores to explain the permeability to water and small water-soluble molecules.

Stratum Corneum—Although the stratum corneum is not a membrane in the same sense as a cell membrane, it offers a barrier to diffusion, which is of significance in the topical application of drugs. The stratum corneum consists of several layers of dead keratinized cutaneous epithelial cells enmeshed in a matrix of keratin fibers and bound together with cementing desmosomes and penetrating tonofibrils of keratin. Varying amounts of lipids and fatty acids from dying cells, sebum and sweat are contained among the dead squamous cells. Immediately beneath the layer of dead cells and above the viable epidermal epithelial cells is a layer of keratohyaline granules and various water-soluble substances, such as alpha-amino acids, purines, monosaccharides and urea.

Both the upper and lower layers of the stratum corneum are involved in the cutaneous barrier to penetration. The barrier to penetration from the surface is in the upper layers for water-soluble substances and the lower layers for lipid-soluble substances, and the barrier to the outward movement of water is in the lowest layer.

Membrane Potentials—Across the cell membrane there exists an electrical potential, always negative on the inside and positive on the outside. If a cell did not have special-membrane electrolyte-transport processes, its membrane potential would be mainly the result of the Donnan equilibrium (see Chapter 14) consequent to the semipermeability of the membrane. Such potentials generally lie between 2 and 5 mv.

A cell with a membrane across which diffusible electrolyte distribution is purely passive would be expected to have a high internal concentration of sodium, such as is true for the erythrocytes of some species. However, the interior of most cells is high in potassium and low in sodium, as depicted in Fig 35-8. This unequal distribution of cations attests to special electrolyte-transport processes and to differential permeabilities of diffusible ions, so that the membrane potential is higher than that which would result from a purely passive Donnan distribution. In nerve tissue or skeletal and cardiac muscle, the membrane potential ranges upwards to about 90 mv. The electrical gradient is on the order of 50,000 v/cm, because of the extreme thinness of the membrane. Obviously, such an intense potential gradient will influence strongly the transmembrane passages of charged drug molecules.

Diffusion and Transport

Transport is the movement of a drug from one place to another within the body. The drug may diffuse freely in uncombined form with a kinetic energy appropriate to its thermal environment, or it may move in combination with extracellular or cellular constituents, sometimes in connection with energy-yielding processes that allow the molecule or complex to overcome barriers to simple diffusion.

Simple Nonionic Diffusion and Passive Transport—Molecules in solution move in a purely random fashion, provided they are not charged and moving in an electrical gradient. Such random movement is called *diffusion;* if the molecule is uncharged, it is called *nonionic diffusion.*

In a population of drug molecules, the probability that during unit time any drug molecule will move across a boundary is directly proportional to the number of molecules adjoining that boundary and, therefore, to the drug concentration. Except at dilutions so extreme that only a few molecules are present, the actual rate of movement (molecules/unit time) is directly proportional to the probability and, therefore, to the concentration. Once molecules have passed through the boundary to the opposite side, their random motion may cause some to return and others to continue to move further away from the boundary. The rate of return is likewise proportional to the concentration on the opposite side of the boundary. It follows that, although molecules are moving in both directions, there will he a net movement from the region of higher to that of lower concentration, and the net transfer will be proportional to the concentration differential. If the boundary is a membrane, which has both substance and dimension, the rate of movement is also directly proportional to the permeability and inversely proportional to the thickness. These factors combine into Fick's Law of Diffusion,

$$\frac{dQ}{dt} = \frac{\overline{D}A(C_1 - C_2)}{x} \qquad (5)$$

where Q is the net quantity of drug transferred across the membrane, t is time, C_1 is the concentration on one side and C_2 on the other, x is the thickness of the membrane, A is the area and $\overline{D}$ is the diffusion coefficient, related to permeability. The equation is more nearly correct if chemical activities are used instead of concentrations. Since a biological membrane is patchy, with pores of different sizes and probably with varying thickness and composition, both $\overline{D}$ and x probably vary from spot to spot. Nevertheless, some mean values can be assumed.

It is customary to combine the membrane factors into a single constant, called a permeability constant or coefficient, P, so that $P = \overline{D}/x$, A in Eq 5 having unit value. The rate of net transport (diffusion) across the membrane then becomes

$$\frac{dQ}{dt} = P(C_1 - C_2) \quad (6)$$

As diffusion continues, C_1 approaches C_2, and the net rate, dQ/dt, approaches zero in exponential fashion characteristic of a first-order process. Equilibrium is defined as that state in which $C_1 = C_2$. The equilibrium is, of course, dynamic, with equal numbers of molecules being transported in each direction during unit time. If water is also moving through the membrane, it may either facilitate the movement of drug or impede it, according to the relative directions of movement of water and drug; this effect of water movement is called *solvent drag.*

Ionic or Electrochemical Diffusion—If a drug is ionized, the transport properties are modified. The probability of penetrating the membrane is still a function of concentration, but it is also a function of the potential difference or electrical gradient across the membrane. A cationic drug molecule will be repelled from the positive charge on the outside of the membrane, and only those molecules with a high kinetic energy will pass through the ion barrier. If the cation is polyvalent, it may not penetrate at all.

Once inside the membrane, a cation simultaneously will be attracted to the negative charge on the intracellular surface of the membrane and repelled by the outer surface; it is said to be moving along the *electrical gradient.* If it also is moving from a higher towards a lower concentration, it is said to be moving along its *electrochemical gradient*, which is the sum of the influences of the electrical field and the concentration differential across the membrane.

Once inside the cell, cations will tend to be kept inside by the attractive negative charge on the interior of the cell, and the intracellular concentration of drug will increase until, by sheer numbers of accumulated drug particles, the outward diffusion or mass escape rate equals the inward transport rate, and electrochemical equilibrium is said to have occurred. At electrochemical equilibrium at body temperature (37°C), ionized drug molecules will be distributed according to the Nernst equation,

$$\pm\log\frac{C_o}{C_i} = \frac{ZE}{61} \quad (7)$$

where C_o is the molar extracellular and C_i the intracellular concentration, Z is the number of charges per molecule and E is the membrane potential in millivolts. Log C_o/C_i is positive when the molecule is negatively charged and negative when the molecule is positively charged.

Facilitated Diffusion—Sometimes a substance moves more rapidly through a biological membrane than can be accounted for by the process of simple diffusion. This accelerated movement is termed *facilitated diffusion.* It is thought to be due to the presence of a special molecule within the membrane, called a *carrier*, with which the transported substance combines. There is considered to be a greater permeability to the carrier–drug complex than to the drug alone, so that the transport rate is enhanced. After the complex traverses the membrane, it dissociates. The carrier must either return to the original side of the membrane to be reused or constantly be produced on one side and eliminated on the other in order for the carrier process to be continuous. Many characteristics of facilitated diffusion, formerly attributed to ion carriers, can be explained by ion exchange. Although facilitated diffusion resembles active transport, below, in its dependence upon a continuous source of energy, it differs in that facilitated diffusion will only transport a molecule along its electrochemical gradient.

Active Transport—Active transport may be defined as energy-dependent movement of a substance through a biological membrane against an electrochemical gradient. It is characterized by the following:

1. The substance is transported from a region of lower to one of higher electrochemical activity.
2. Metabolic poisons interfere with transport.
3. The transport rate approaches an asymptote (ie, saturates) as concentration increases.
4. The transport system usually shows a requirement for specific chemical structures.
5. Closely related chemicals are competitive for the transport system.

Many drugs are secreted from the renal tubules into urine, from liver cells into bile or from the cerebrospinal fluid into blood by active transport, but the role of active transport of drugs in the distribution into most body compartments and tissues is less well known. Active transport is required for the penetrance of a number of sympathomimetics into neural tissue and for the movement of several anticancer drugs across cell membranes.

Pinocytosis and Exocytosis—Many, perhaps all, cells are capable of a type of phagocytosis called *pinocytosis.* The cell membrane has been observed to invaginate into a saccular structure containing extracellular materials and then pinch off the saccule at the membrane, so that the saccule remains as a vesicle or vacuole within the interior of the cell. Since metaholic activity is required and since an extracellular substance may be transported against an electrochemical gradient, pinocytosis shows some of the same characteristics as active transport. However, pinocytosis is relatively slow and inefficient compared to most active transport, except in gastrointestinal absorption, in which pinocytosis is of considerable importance.

It is not known to what extent pinocytosis contributes to the transport of most drugs, but many macromolecules and even larger particles can be absorbed by the gut. Pinocytosis probably explains the oral efficacy of the Sabin polio vaccine. Some drugs themselves affect pinocytosis; eg, adrenal glucocorticoids markedly inhibit the process in macrophages and other cells involved in inflammation.

Exocytosis is more or less the reverse of pinocytosis. Granules, vacuoles or other organelles within the cell move to the cell membrane, fuse with it and extrude their contents into the interstitial space.

Physicochemical Factors in Penetrance

Drugs and other substances may traverse the membrane primarily either through the pores or by dissociation into the membrane lipids and subsequent diffusion from the membrane into the cytosol or other fluid on the far side of the membrane. The physicochemical prerequisites are different according to which route is taken. To pass through the pores, the "diameter" of the molecule must be smaller than the pore, but the molecule can be longer than the pore diameter. The probability that a long, thin molecule will be oriented properly is low, unless there is also bulk flow and the transmembrane passage of large molecules is slow.

Water-soluble molecules with low lipid solubility usually are thought to pass through the membrane mainly via the pores and, to a small extent, by pinocytosis, but recent work with lipid monolayers suggests that small water-soluble molecules also may he able to pass readily through the lipid, and the necessity of postulating the existence of pores has been questioned. Nevertheless, experimental data on penetrance overwhelmingly favor the concept of passage of water-soluble lipid-insoluble substances through pores. If there is a membrane carrier or active transport system, a low solubility of the drug in membrane lipids is no impediment to penetration, since the drug-carrier complex is assumed to have an appropriate solubility and energy from an active transport system enables the drug to penetrate the energy barrier "imposed by the lipids." Actually, the lipids are not an important energy barrier; rather, the barrier is the force

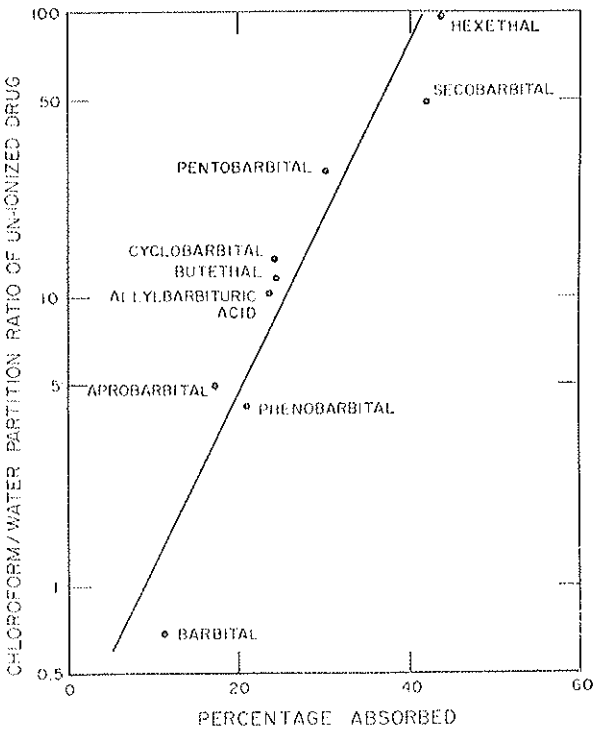


Fig 35-9. The relationship of absorption of the un-ionized forms of drugs from the colon of the rat to the chloroform:water partition coefficient (courtesy, Schanker[6]).

of attraction of the solvent water for its dipolar to polar solute, so that it is difficult for the solute to leave the water and enter the lipid.

Drugs with a high solubility in the membrane lipids, of course, pass easily through the membrane. Even when their dimensions are small enough to permit passage through pores, lipid-soluble drugs primarily pass through the membrane lipids, not only because chemical partition favors the lipid phase but also because the surface area occupied by pores is only a small fraction of the total membrane area.

Lipid Solubility and Partition Coefficients—As early as 1902, Overton investigated the importance of lipid solubility to the penetrance and absorption of drugs. Eventually, it was recognized that more important than lipid solubility was the lipid–water distribution coefficient; ie, a high lipid solubility does not favor penetrance unless the water solubility is low enough so that the drug is not entrained in the aqueous phase.

In Fig 35-9 is illustrated the relationship between the chloroform–water partition coefficient and the colonic absorption of barbiturates. Chloroform probably is not the optimal lipid solvent for such a study, and natural lipids from nerve or other tissues have been shown to be superior in the few instances in which they have been employed. Nevertheless, the correlation shown in the figure is a convincing one.

When the water solubility of a substance is so low that a significant concentration in water or extracellular fluid cannot be achieved, absorption may be negligible in spite of a favorable partition coefficient. Hence, mineral oil, petrolatum, etc are virtually unabsorbed. The optimal partition coefficient for permeation of the skin appears to be lower than that for the permeation of the cell membrane, perhaps being as low as one.

Dipolarity, Polarity and Nonionic Diffusion—The partition coefficient of a drug depends upon the polarity and the size of the molecule. Drugs with a high dipole moment, even though un-ionized, have a low lipid solubility and, hence, poor penetrance. An example of a highly dipolar substance with a low partition coefficient, which does not penetrate into cells, is sulfisoxazole. Sulfadiazine is somewhat less dipolar, has a chloroform–water partition coefficient ten times that of sulfisoxazole and readily penetrates cells. Ionization not only diminishes lipid solubility greatly but also may impede passage through charged membranes (see *Ionic Diffusion*, page 709).

It is often stated that ionized molecules do not penetrate membranes, except for ions of small diameter. This is not necessarily true, because of the presence of membrane carriers for some ions, which effectively may shield or neutralize the charge (ion-pair formation). The renal tubular transport systems, which transport such obligate ions as tetraethylammonium, probably form ion-pairs. Furthermore, if an ionized molecule has a large nonpolar moiety such that an appreciable lipid solubility is imparted to the molecule in spite of the charge, the drug may penetrate, although usually at a slow rate. For example, various morphinan derivatives are absorbed passively from the stomach even though they

Table I—Rates of Entry of Drugs in CSF and the Degrees of Ionization of Drugs at pH 7.4[7]

Drug	% binding to plasma protein	pK_a[a]	% un-ionized at pH 7.4	Permeability constant (P min^{-1}) ± S.E.
		Drugs mainly ionized at pH 7.4		
5-Sulfosalicylic acid	22	(strong)	0	<0.0001
N-Methylnicotinamide	<10	(strong)	0	0.0005 ± 0.00006
5-Nitrosalicylic acid	42	2.3	0.001	0.001 ± 0.0001
Salicylic acid	40	3.0	0.004	0.006 ± 0.0004
Mecamylamine	20	11.2	0.016	0.021 ± 0.0016
Quinine	76	8.4	9.09	0.078 ± 0.0061
		Drugs mainly un-ionized at pH 7.4		
Barbital	<2	7.5	55.7	0.026 ± 0.0022
Thiopental	75	7.6	61.3	0.50 ± 0.051
Pentobarbital	40	8.1	83.4	0.17 ± 0.014
Aminopyrine	20	5.0	99.6	0.25 ± 0.020
Aniline	15	4.6	99.8	0.40 ± 0.042
Sulfaguanidine	6	>10.0[b]	>99.8	0.003 ± 0.0002
Antipyrine	8	1.4	>99.9	0.12 ± 0.013
N-Acetyl-4-aminoantipyrine	<3	0.5	>99.9	0.012 ± 0.0010

[a] The dissociation constant of both acids and bases is expressed as a pK_a—a negative logarithm of the acidic dissociation constant.

[b] Sulfaguanidine has a very weakly acidic group (pK_a > 10) and two very weakly basic groups (pK_a 2.75 and 0.5). Consequently, the compound is almost completely undissociated at pH 7.4.

are ionized completely at the pH of gastric fluid. Nevertheless, when a drug is a weak acid or base, the un-ionized form, with a favorable partition coefficient, passes through a biological membrane so much more readily than the ionized form that, for all practical purposes, only the un-ionized form is said to pass through the membrane. This has become known as the *principle of nonionic diffusion*.

This principle is the reason that only the concentrations of the un-ionized form of the barbiturates are plotted in Fig 35-9.

For the purpose of further illustrating the principle, Table I is provided. In the table, the permeability constants for penetrance into the cerebral spinal fluid of rats are higher for un-ionized drugs than for ionized ones. The apparent exceptions—barbital, sulfaguanidine and acetylaminoantipyrine—may be explained by the dipolarity of the un-ionized molecules. With barbital, the two lipophilic ethyl groups are too small to compensate for the considerable dipolarity of the un-ionized barbituric acid ring; also it may be seen that barbital is appreciably ionized, which contributes to the relatively small permeability constant. Sulfaguanidine and acetylaminoantipyrine are both very polar molecules. Mecamylamine also might be considered an exception, since it shows a modest permeability even though strongly ionized; there is no dipolarity in mecamylamine except in the amino group.

Absorption of Drugs

Absorption is the process of movement of a drug from the site of application into the extracellular compartment of the body. Inasmuch as there is a great similarity among the various membranes that a drug may pass through in order to gain access to the extracellular fluid, it might be expected that the particular site of application (or *route*) would make little difference to the successful absorption of the drug. In actual fact, it makes a great deal of difference; many factors, other than the structure and composition of the membrane, determine the ease with which a drug is absorbed. These factors are discussed in the following sections, along with an account of the ways that drug formulations may be manipulated to alter the ability of a drug to be absorbed readily.

Routes of Administration

Drugs may be administered by many different routes. The various routes include oral, rectal, sublingual or buccal, parenteral, inhalation and topical. The choice of a route depends upon both convenience and necessity.

Oral Route—This is obviously the most convenient route for access to the systemic circulation, providing that various factors do not militate against this route. Oral administration does not always give rise to sufficiently high plasma concentrations to be effective; some drugs are absorbed unpredictably or erratically; patients occasionally have an absorption malfunction. Drugs may not be given by mouth to patients with gastrointestinal intolerance, or who are in preparation for anesthesia or who have had gastrointestinal surgery. Oral administration also is precluded in coma.

Rectal Route—Drugs that ordinarily are administered by the oral route usually can be administered by injection or by the alternative *lower enteral* route, through the anal portal into the rectum or lower intestine. With regard to the latter, *rectal suppositories* or *retention enemas* formerly were used quite frequently, but their popularity has abated somewhat, owing to improvements in parenteral preparations. Nevertheless, they continue to be valid and, sometimes, very important ways of administering a drug, especially in pediatrics and geriatrics. In Fig 35-10 the availability of a drug by retention enema may be compared with that by the intravenous and oral route and rectal suppository administration. It is apparent that the retention enema may be a very satisfactory means of administration but that rectal suppositories may be inadequate where rapid absorption and high plasma levels are required. The illustration is not intended to lead the reader to the conclusion that a retention enema always will give more prompt and higher blood levels than the oral route, for converse findings for the same drug have been reported,[21] but, rather, to show that the retention enema may offer a useful substitute for the oral route.

Sublingual or Buccal Route—Even though an adequate plasma concentration eventually may be achievable by the oral route, it may rise much too slowly for use in some situations where a rapid response is desired. In such situations parenteral therapy usually is indicated. However, the patients with angina pectoris may get quite prompt relief from an acute attack by the *sublingual* or *buccal* administration of nitroglycerin, so that parenteral administration may be avoided. When only small amounts of drugs are required to gain access to the blood, the buccal route may be very satisfactory, providing the physicochemical prerequisites for absorption by this route are present in the drug and dosage form. Only a few drugs may be given successfully by this route.

Parenteral Routes—These routes, by definition, include any route other than the oral-gastrointestinal (enteral) tract, but in common medical usage the term excludes topical administration and includes only various hypodermic routes. Parenteral administration includes the intravenous, intramuscular and subcutaneous routes. Parenteral routes may be employed whenever enteral routes are contraindicated (see above) or inadequate.

The *intravenous* route may be preferred on occasion, even when a drug may be well-absorbed by the oral route. There is no delay imposed by absorption before the administered drug reaches the circulation, and blood levels rise virtually

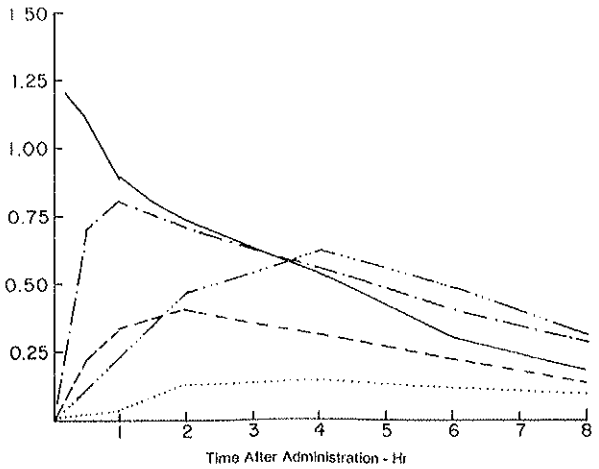


Fig 35-10. Blood concentration in mg/100 mL of theophylline (ordinates) following administration to humans of aminophylline in the amounts and by the routes indicated. Doses: per 70 kg. Theophylline-ethylenediamine by various routes: —— *intravenous, 0.5 g;* – • – • – *retention enema, 0.5 g;* — ··· — ··· — *oral tablets-Pl., 0.5 g;* - - - *oral tablets-Pl., 0.3 g;* ···· *rectal suppository, 0.5 g* (courtesy, Truitt, *et al*,[20] adapted).

as rapidly as the time necessary to empty the syringe or infusion bottle. Consequently, the intravenous route is the preferred route when an emergency calls for an immediate response.

In addition to the rapid rise in plasma concentration of drug, another advantage of intravenous administration is the greater predictability of the peak plasma concentration, which, with some drugs, can be calculated with a fair degree of precision. Smaller doses generally are required by the intravenous than by other routes, but this usually affords no advantage, inasmuch as the sterile injectable dosage form costs more than enteric preparations and the requirements for medical or paramedical supervision of administration also may add to the cost and inconvenience.

Because of the rapidity with which drug enters the circulation, dangerous side effects to the drug may occur which are often not extant by other routes. The principal untoward effect is a depression of cardiovascular function, which is often called *drug shock*. Consequently, some drugs must be given quite slowly to avoid vasculotoxic concentrations of drug in the plasma. Acute, serious, allergic responses also are more likely to occur by the intravenous route than by other routes.

Many drugs are too irritant to be given by the oral, intramuscular or subcutaneous route and must, of necessity, be given intravenously. However, such drugs also may cause damage to the veins (phlebitis) or, if extravasated, cause necrosis (slough) around the injection site. Consequently, such irritant drugs may be diluted in isotonic solutions of saline, dextrose or other media and given by slow infusion, providing that the slower rate of delivery does not negate the purpose of the administration in emergency situations.

Absorption by the *intramuscular* route is relatively fast and this parenteral route may be used where an immediate effect is not required but a prompt effect is desirable. Intramuscular deposition also may be made of certain repository preparations, rapid absorption not being desired. Absorption from an intramuscular depot is more predictable and uniform than from a subcutaneous site.

Irritation around the injection site is a frequent accompaniment of intramuscular injection, depending upon the drug and other ingredients. Because of the dangers of accidental intravenous injection, medical supervision generally is required. Sterilization is necessary.

In *subcutaneous* administration the drug is injected into the alveolar connective tissue just below the skin. Absorption is slower than by the intramuscular route but, nevertheless, may be prompt with many drugs. Often, however, absorption by this route may be no faster than by the oral route. Therefore, when a fairly prompt response is desired with some drugs, the subcutaneous route may not offer much advantage over the oral route, unless for some reason the drug cannot be given orally.

The slower rate of absorption by the subcutaneous route is usually the reason why the route is chosen, and the drugs given by this route are usually those in which it is desired to spread the action out over a number of hours, in order to avoid either too intense a response, too short a response or frequent injections. Examples of drugs given by this route are insulin and sodium heparin, neither of which is absorbed orally and both of which should be absorbed slowly over many hours. In the treatment of asthma, epinephrine usually is given subcutaneously to avoid the dangers of rapid absorption and consequent dangerous cardiovascular effects. Many repository preparations, including tablets or pellets, are given subcutaneously. As with other parenteral routes, irritation may occur. Sterile preparations also are required. However, medical supervision is not required always and self-administration by this route is customary with certain drugs, such as insulin.

Intradermal injection, in which the drug is injected into, rather than below the dermis, is rarely employed, except in certain diagnostic and test procedures, such as screening for allergic or local irritant responses.

Occasionally, even by the intravenous route, it is not possible, practical or safe to achieve plasma concentrations high enough so that an adequate amount of drug penetrates into special compartments, such as the cerebrospinal fluid, or various cavities, such as the pleural cavity. The brain is especially difficult to penetrate with water-soluble drugs. The name *blood-brain barrier* is applied to the impediment to penetration. When drugs do penetrate, the choroid plexus often secretes them back into the blood very rapidly, so that adequate levels of drugs in the cerebrospinal fluid may be difficult to achieve. Consequently, *intrathecal** or *intraventricular* administration may be indicated.

Body cavities such as the pleural cavity normally are wetted by a small amount of effusate which is in diffusion equilibrium with the blood and, hence, is accessible to drugs. However, infections and inflammations may cause the cavity to fill with serofibrinous exudate which is too large to be in rapid diffusion equilibrium with the blood. *Intracavitary* administration, thus, may be required. It is extremely important that sterile and nonirritating preparations be used for intrathecal or intracavitary administration.

Inhalation Route—Inhalation may be employed for delivering gaseous or volatile substances into the systemic circulation, as with most general anesthetics. Absorption is virtually as rapid as the drug can be delivered into the alveoli of the lungs, since the alveolar and vascular epithelial membranes are quite permeable, blood flow is abundant and there is a very large surface for absorption.

Aerosols of nonvolatile substances also may be administered by inhalation, but the route is used infrequently for delivery into the systemic circulation because of various factors which contribute to erratic or difficult-to-achieve blood levels. Whether or not an aerosol reaches and is retained in pulmonary alveoli depends critically upon particle size. Particles greater than 1 μm in diameter tend to settle in the bronchioles and bronchi, whereas particles less than 0.5 μm fail to settle and mainly are exhaled. Aerosols are employed mostly when the purpose of administration is an action of the drug upon the respiratory tract itself. An example of a drug commonly given as an aerosol is isoproterenol, which is employed to relax the bronchioles during an asthma attack.

Topical Route—Topical administration is employed to deliver a drug at, or immediately beneath, the point of application. Although occasionally enough drug is absorbed into the systemic circulation to cause systemic effects, absorption is too erratic for the topical route to be used routinely for systemic therapy. However, various transdermal preparations of nitroglycerin and clonidine are employed quite successfully for systemic use. Some work with aprotic solvent vehicles such as dimethyl sulfoxide (DMSO) also has generated interest in topical administration for systemic effects. A large number of topical medicaments are applied to the skin, although topical drugs are also applied to the eye, nose and throat, ear, vagina, etc.

In man, percutaneous absorption probably occurs mainly from the surface. Absorption through the hair follicles occurs, but the follicles in man occupy too small a portion of the total integument to be of primary importance. Absorption through sweat and sebaceous glands generally appears to be minor. When the medicament is rubbed on vigorously,

* Intrathecal administration denotes administration into the cerebrospinal fluid at any level of the cerebrospinal axis, including injection into the cerebral ventricles, which is the most common mode of intrathecal administration.

the amount of the preparation that is forced into the hair follicles and glands is increased. Rubbing also forces some material through the stratum corneum without molecular dispersion and diffusion through the barrier. Rather large particles of substances such as sulfur have been demonstrated to pass intact through the stratum corneum. When the skin is diseased or abraded, the cutaneous barrier may be disrupted or defective, so that percutaneous absorption may be increased. Since much of a drug that is absorbed through the epidermis diffuses into the circulation without reaching a high concentration in some portions of the dermis, systemic administration may be preferred in lieu of, or in addition to, topical administration.

Factors That Affect Absorption

In addition to the physicochemical properties of drug molecules and biological membranes, various factors affect the rate of absorption and determine, in part, the choice of route of administration.

Concentration—It is self-evident that the concentration, or, more exactly, the thermodynamic activity, of a drug in a drug preparation will have an important bearing upon the rate of absorption, since the rate of diffusion of a drug away from the site of administration is directly proportional to the concentration. Thus, a 2% solution of lidocaine will induce local anesthesia more rapidly than a 0.2% solution. However, drugs administered in solid form are not absorbed necessarily at the maximal rate (see *Physical State of Formulation and Dissolution Rate*, below).

After oral administration the concentration of drugs in the gut is a function of the dose, but the relationship is not necessarily linear. Drugs with a low aqueous solubility (eg, digitoxin) quickly saturate the gastrointestinal fluids, so that the rate of absorption tends to reach a limit as the dose is increased. The peptizing and solubilizing effects of bile and other constituents of the gastrointestinal contents assist in increasing the rate of absorption but are in themselves somewhat erratic. Furthermore, many drugs affect the rates of gastric, biliary and small intestinal secretion, which causes further deviations from a linear relationship between concentration and dose.

Drugs that are administered subcutaneously or intramuscularly also may not show always a direct linear relationship between the rate of absorption and the concentration of drug in the applied solution, because osmotic effects may cause dilution or concentration of the drug, if the movement of water or electrolytes is different from that of the drug. Whenever possible, drugs for hypodermic injection are prepared as isotonic solutions. Some drugs affect the local blood flow and capillary permeability, so that at the site of injection there may be a complex relationship of concentration achieved to the concentration administered.

Physical State of Formulation and Dissolution Rate—The rate of absorption of a drug may be affected greatly by the rate at which the drug is made available to the biological fluid at the site of administration. The intrinsic physicochemical properties, such as solubility and the thermodynamics of dissolution, are only some of the factors which affect the rate of dissolution of a drug from a solid form. Other factors include not only the unavoidable interactions among the various ingredients in a given formulation but also deliberate interventions to facilitate dispersion (eg, comminution, Chapter 75 and dissolution, Chapter 31) or retard it (eg, coatings, Chapter 90 and slow-release formulations, Chapter 91). There are also factors that affect the rate of delivery from liquid forms. For example, a drug in a highly viscous vehicle is absorbed more slowly from the vehicle than a drug in a vehicle of low viscosity; in oil-in-water emulsions, the rate depends upon the partition coefficient. These manipulations are the subject of biopharmaceutics (see Chapter 91).

Area of Absorbing Surface—The area of absorbing surface is an important determinant of the rate of absorption. To the extent that the therapist must work with the absorbing surfaces available in the body, the absorbing surface is not subject to manipulation. However, the extent to which the existing surfaces may be used is subject to variation. In those rare instances in which percutaneous absorption is intended for systemic administration, the entire skin surface is available.

Subsequent to subcutaneous or intramuscular injections, the site of application may be massaged in order to spread the injected fluid from a compact mass to a well-dispersed deposit. Alternatively, the dose may be divided into multiple small injections, although this recourse is generally undesirable.

The different areas for absorption afforded by the various routes account, in part, for differences in the rates of absorption by those routes. The large alveolar surface of the lungs allows for extremely rapid absorption of gases, vapors and properly aerosolized solutions; with some drugs the rate of absorption may be nearly as fast as intravenous injection. In the gut the small intestine is the site of the fastest, and hence most, absorption because of the small lumen and highly developed villi and microvilli; the stomach has a relatively small surface area, so that even most weak acids are absorbed predominately in the small intestine despite a pH partition factor that should favor absorption from the stomach (see *The pH Partition Principle*, page 716).

Vascularity and Blood Flow—Although the thermal velocity of a freely diffusible average drug molecule is on the order of meters per second, in solution the rate at which it will diffuse away from a reference point will be much slower. Collisions with water and/or other molecules, which cause a random motion, and the forces of attraction between the drug and water or other molecules slow the net mean velocity.

The time taken to traverse a given distance is a function of the square of the distance; on the average it would take about 0.01 sec for a net outward movement of 1 μm, 1 sec for 10 μm, 100 sec for 100 μm, etc. In a highly vascular tissue, such as skeletal muscle, in which there may be more than 1000 capillaries/sq mm of cross section, a drug molecule would not have to travel more than a few microns, hence, less than a second on the average, to reach a capillary from a point of extravascular injection.

Once the drug reaches the blood, diffusion is not important to transport and the rate of blood flow determines the movement. The velocity of blood flow in a capillary is about 1 mm/sec, which is 100 times faster than the mean net velocity of drug molecules 1 mm away from their injection site. The velocity of blood flow is even faster in the larger vessels. Less than a minute is required to distribute drug molecules from the capillaries at the injection site to the rest of the body.

From the above discussion it follows that absorption is most rapid in the vascular tissues. Drugs are absorbed more rapidly from intramuscular sites than from less vascular subcutaneous sites, etc. Despite the small absorbing surface for buccal or sublingual absorption, the high vascularity of the buccal, gingival and sublingual surfaces favors an unexpectedly high rate of absorption. Because of hyperemia, absorption will be faster from inflamed than from normal areas, unless the presence of edema lengthens the mean distance between capillaries and, thus, negates the effects of hyperemia on absorption.

Vasoconstriction may have a profound effect upon the rate of absorption. When a local effect of a drug is desired,

as in local anesthesia, absorption away from the infiltered site may be impeded greatly by vasoconstrictors included in the preparation. Unwanted vasoconstriction sometimes may cause serious problems. For example, in World War II many wounded soldiers were given subcutaneous morphine without evident effect. As a result, injections were sometimes repeated more than once. When the patient was removed to the field hospital, toxic effects would occur suddenly. The explanation is that cold-induced vasoconstriction occurred in the field; when the patient was warmed in the hospital, vasodilation would result and the victim would be flooded with drug. Shock also contributes to the effect, since during shock the blood flow is diminished and there also may be a superimposed vasoconstriction; repair of the shock condition then facilitates absorption.

Extravascularly injected molecules too large to pass through the capillary endothelium will, of necessity, enter the systemic circulation through the lymph. Thus, the lymph flow may be important to the absorption of a few drugs.

Movement—A number of factors combine so that movement at the site of injection increases the rate of absorption. In the intestine, segmental movements and peristalsis aid in dividing and dispersing the drug mass. The continual mixing of the chyme helps keep the concentration maximal at the mucosal surface. The pressures developed during segmentation and peristalsis also may favor a small amount of filtration. Movement at the site of hypodermic injection also favors absorption, since it tends to force the injected material through the tissue, increasing the surface area of drug mass and decreasing the mean distance to the capillaries. Movement also increases the flow of blood and lymph. The selection of a site for intramuscular injection may be determined by the amount of expected movement, according to whether the preparation is intended as a fast-acting or a repository preparation.

Gastric Motility and Emptying—The motility of the stomach is more important to the rate at which an orally administered drug is passed on to the small intestine than it is to the rate of absorption from the stomach itself, since for various reasons noted above, absorption from the stomach is usually of minor importance.

The average emptying time of the unloaded stomach is about 40 min and the half-time is around 10 min, though it varies according to its contents, reflex and psychological factors, as well as the action of certain autonomic drugs or disease. The effect of food to delay absorption is due, in part, to its action to prolong emptying time. The emptying time causes a delay in the absorption of drug, which may be unfavorable or favorable according to what is desired. In the case of therapy with antacids, gastric emptying is a nuisance, since it removes the antacid from the stomach where it is needed.

Solubility and Binding—The dissolution of drugs of low solubility is generally a slow process. Indeed, low solubility is the result of a low rate of departure of drug molecules from the undispersed phase. Furthermore, since the concentration around the drug mass is low, the concentration gradient from the site of deposition to the plasma is small and the rate of diffusion is low, accordingly.

When it is desired that a drug have a prolonged action but not a high plasma concentration, a derivative of low solubility is often sought. The "insoluble" estolates and other esters of several steroids have durations of action of weeks because of the slow rates of absorption from the sites of injection. Insoluble salts or complexes of acidic or basic drugs also are employed as repository preparations; for example, the procaine salt of penicillin G has a low solubility and is used in a slow release form of the antibiotic.

The solubility of certain macromolecules is dependent critically on the ionization of substituent groups. When they are amphiprotic, they are least soluble at their isoelectric pH. Insulin is normally soluble at the pH of the extracellular fluid, but by combining insulin with the right proportion of a basic protein, such as protamine, the isoelectric pH can be made to be approximately 7.4 and the complex can be used as a low-solubility prolonged-action drug. For more details, see Chapter 91.

Some drugs may bind with natural substances at or near the site of application. The strongly ionized mucopolysaccharides in connective tissue, ground substance and mucous secretions of the gut are retardants to the absorption of a number of drugs, especially large cationic or polycationic molecules. In the gut, the binding is the least at low pH, which should favor absorption of large cations from the stomach; however, absorption from the stomach is slow (see above), so that the absorption of large cations occurs mainly in the upper duodenum where the pH is still relatively low. Pharmacologically inactive quaternary ammonium compounds sometimes are included in an oral preparation of a quaternary ammonium drug for the purpose of saturating the binding sites of mucin and other mucopolysaccharides and, thereby, enhancing the absorption of drug.

In addition to mucopolysaccharides in mucous secretions, food in the gastrointestinal track binds many drugs and slows absorption. Antacids, especially aluminum hydroxide plus other basic aluminum compounds and magnesium trisilicate, bind amine and ammonium drugs and interfere with absorption.

Donnan Effect—The presence of a charged macromolecule on one side of a semipermeable membrane (impermeable to the macromolecule) will alter the concentration of permeant ionized particles according to the Donnan equilibrium (page 716). Accordingly, drug molecules of the same charge as the macromolecule will be constrained to the opposite side of the membrane. The presence of appropriately charged macromolecules not only will influence the distribution of drug ions in accordance with the Donnan equation but also increase the rate of transfer of the drug across the membrane, because of mutual ionic repulsion. This effect is sometimes used to facilitate the absorption of ionizable drugs from the gastrointestinal tract. The Donnan effect also operates to retard the absorption of drug ions of opposite charge; however, the mutual electrostatic attraction of a macromolecule and drug ion generally results in actual binding, which is more important than the Donnan effect.

Vehicles and Absorption Adjuvants—Drugs that are to be applied topically to the skin and mucous membranes often are dissolved in vehicles that are thought to enhance penetrance. For a long time it was thought that oleaginous vehicles promoted the absorption of lipid-soluble drugs. However, the role and effect of the vehicle has proven to be quite complex. In the skin at least five factors are involved:

1. The effect of the vehicle to alter the hydration of the keratin in the barrier layer.
2. The effect of the vehicle to promote or prevent the collection of sweat at the surface of the skin.
3. The partition coefficient of the drug in a vehicle–water system.
4. The permeability of the skin to the undissolved drug.
5. The permeability of the skin to the vehicle.

The effect of the vehicle to aid in the access of the drug to the hair follicles and sebaceous glands also may be involved, although in man the follicles and glands are probably ordinarily of minor importance to absorption.

A layer of oleaginous material over the skin prevents the evaporation of water, so that the stratum corneum may become macerated and more permeable to drugs. In dermatology it is sometimes the practice to wrap the site of application with plastic wrap or some other waterproof material for the purpose of increasing the maceration of the stratum

corneum. However, the layer of perspiration that forms under an occlusive vehicle may become a barrier to the movement of lipid-soluble drugs from the vehicle to the skin, but it may facilitate the movement of water-soluble drugs. Conversely, polyethylene glycol vebicles remove the perspiration and dehydrate the barrier, which decreases the permeability to drugs; sucb vehicles remove the aqueous medium through which water-soluble drugs may pass down into the stratum corneum but at the same time facilitate the transfer of lipid-soluble drugs from the vehicle to the skin.

Even in the absence of a vehicle, it is not clear what physicochemical properties of a drug favor cutaneous penetration, high lipid-solubility being a prerequisite, according to some authorities, and an ether–water partition coefficient of approximately one, according to others. Yet, the penetration of ethanol and dibromomethane are nearly equal, and other such enigmas exist. It is not surprising, then, that the effects of vehicles are not altogether predictable.

A general statement might be made that if a drug is quite soluble in a poorly absorbed vehicle, the vehicle will retard the movement of the drug into the skin. For example, salicylic acid is 100 times as permeant when absorbed from water than from polyethylene glycol and pentanol is 5 times as permeant from water as from olive oil. Yet, ethanol penetrates 5 times faster from olive oil than from either water or ethanol, all of which denies the trustworthiness of generalizations about vehicles.

Since the 1960s, there has been much interest in certain highly dielectric aprotic solvents, especially dimethyl sulfoxide (DMSO). Such substances generally prove to be excellent solvents for both water- and lipid-soluble compounds and for some compounds not soluble in either water or lipid solvents. The extraordinary solvent properties probably are due to a high polarizability and van der Waals bonding capacity, a high degree of polarization (dipole moment) and a lack of association through hydrogen bonding. As a vehicle, DMSO greatly facilitates the permeation of the skin and other biological membranes by numerous drugs, including such large molecules as insulin. The mechanism is understood poorly. Such vehicles have a potential for many important uses, but they are at present only experimental, pending further investigations on toxicity.

From time to time, a claim is made that a new ingredient of a tablet or elixir enhances the absorption of a drug, and a comparison of plasma levels of the old and new preparations seems to support the claim. Upon further investigation, however, it may be revealed that the new so-called absorption adjuvant is replacing an ingredient that previously bound the drug or delayed its absorption; thus, the new "adjuvant" is not an adjuvant but rather it is only a nondeterrent.

Other Factors—A number of other less-well-defined factors affect the absorption of drugs, some of which may operate, in part, through factors already cited above. Disease or injury has a considerable effect upon absorption. For example, debridement of the stratum corneum increases tbe permeability to topical agents, meningitis increases the permeability of the blood–brain barrier, biliary insufficiency decreases the absorption of lipid-soluble substances from the intestine and acid–base disturbances can affect the absorption of weak acids or bases. Certain drugs, such as ouabain, that affect active transport processes may interfere with the absorption of certain other drugs. The condition of the *ground substance*, or "intracellular cement," probably bears on the absorption of certain types of molecules. Hyaluronidase, which depolymerizes the mucopolysaccharide ground substance, can be demonstrated to facilitate the absorption of some, but not all, drugs from subcutaneous sites.

Drug Disposition

The term *drug disposition* is used here to include all processes which tend to lower the plasma concentration of drug, as opposed to drug absorption, which elevates the plasma level. Consequently, the distribution of drugs to the various tissues will be considered under *Disposition*. Some authors use the term disposition synonymously with elimination, that is, to include only those processes which decrease the amount of drug in the body. In the present context, disposition comprises three categories of processes: distribution, biotransformation and excretion.

Distribution, Biotransformation and Excretion

The term *distribution* is self-explanatory. It denotes the partitioning of a drug among the numerous locations where a drug may be contained within the body. *Biotransformations* are the alterations in the chemical structure of a drug that are imposed upon it by the life processes. *Excretion* is, in a sense, the converse of absorption, namely, the transportation of the drug, or its products, out of the body. The term applies whether or not special organs of excretion are involved.

Distribution

The body may be considered to comprise a number of *compartments:* enteric (gastrointestinal), plasma, interstitial, cerebrospinal fluid, bile, glandular secretions, urine, storage vesicles, cytoplasm or intracellular space, etc. Some of these "compartments," such as urine and secretions, are open-ended, but since their contents relate to those in the closed compartments, they also must be included.

At first thought, it may seem that if a drug were distributed passively (ie, by simple diffusion) and the plasma concentration could be maintained at a steady level, the concentration of a drug in the water in all compartments ought to become equal. It is true that some substances, such as ethanol and antipyrine, are distributed nearly equally throughout the body water, but they are more the exception than the rule. Such substances are mainly small, uncharged, nondissociable, highly water-soluble molecules.

The condition of small size and high water solubility allows for passage through the pores without the necessity of carrier or active transport. Small size also places a limit on van der Waals binding energy and configurational complementariness, so that binding to proteins in plasma, or cells, is slight. The presence of a charge on a drug molecule makes for unequal distribution across charged membranes, in accordance with the Donnan distribution (see below). Dissociability causes unequal distribution when there is a pH differential between compartments, as discussed under *The pH Partition Principle* (see below). Thus, even if a drug is distributed passively, its distribution may be uneven throughout the body. When active transport into, or a rapid metabolic destruction occurs within, some compartments, uneven distribution is also inevitable.

The pH Partition Principle—An important consequence of nonionic diffusion is that a difference in pH between two compartments will have an important influence upon the partitioning of a weakly acidic or basic drug between those compartments. The partition is such that the un-ionized form of the drug has the same concentration in both compartments, since it is the form that is freely diffusible; the ionized form in each compartment will have the

concentration that is determined by the pH in that compartment, the pK and the concentration of the un-ionized form. The governing effect of pH and pK on the partition is known as the *pH partition principle*.

To illustrate the principle, consider the partition of salicylic acid between the gastric juice and the interior of a gastric mucosal cell. Assume the pH of the gastric juice to be 1.0, which it occasionally becomes. The pK_a of salicylic acid is 3.0 (Martin[10] provides one source of pK values of drugs). With the Henderson–Hasselbach equation (see page 242) it may be calculated that the drug is only 1% ionized at pH 1.0.* The intracellular pH of most cells is about 7.0. Assuming the pH of the mucosal cell to be the same, it may be calculated that salicylic acid will be 99.99% ionized within the cells. Since the concentration of the un-ionized form is theoretically the same in both gastric juice and mucosal cells, it follows that the total concentration of the drug (ionized + un-ionized) within the mucosal cell will be 10,000 times greater than in gastric juice. This is illustrated in Fig 35-11. Such a relatively high intracellular concentration can have important osmotic and toxicologic consequences.

Had the drug been a weak base instead of an acid, the high concentration would have been in the gastric juice. In the small intestine, where the pH may range from 7.5 to 8.1, the partition of a weak acid or base will be the reverse of that in the stomach, but the concentration differential will be less, because the pH differential from lumen to mucosal cells, etc, will be less. The reversal of partition as the drug moves from the stomach to the small intestine accounts for the phenomenon that some drugs may be absorbed from one gastrointestinal segment and returned to another. The weak base, atropine, is absorbed from the small intestine, but, because of pH partition, it is "secreted" into the gastric juice.

The pH partition of drugs has never been demonstrated to be as marked as that illustrated in Fig 35-11 and in the text. Not only do many drug ions probably pass through the pores of the membrane to a significant extent, but also some may pass through the lipid phase, as explained above for the morphinans and mecamylamine. Furthermore, ion-pair formation in carrier transport also bypasses nonionic diffusion. All processes that tend toward an equal distribution of drugs across membranes, and among compartments, will cause further deviations from theoretical predictions of pH partition.

Electrochemical and Donnan Distribution—A drug ion may be distributed passively across a membrane in accordance with the membrane potential, the charge on the drug ion and the Donnan effect. The relationship of the membrane potential to the passive distribution of ions is expressed quantitatively by the Nernst equation (Eq 7, page 709) and already has been discussed. Barring active transport, pH partition and binding, the drug will be said to be distributed according to the electrical gradient or to its "equilibrium" potential. If the membrane potential is 90 mv, the concentration of a univalent cation will be 30 times as high within the cell as without; if the drug cation is divalent, the ratio will be 890. The distribution of anions would be just the reverse. If the membrane potential is but 9 mv, the ratio for a univalent cation will be only 1.4 and for a divalent cation only 2.0. It, thus, can be seen how important membrane potential may be to the distribution of ionized drugs.

It was pointed out under *Membrane Potentials*, page 707, that large potentials derive from active transport of ions but that small potentials may result from Donnan distribution. Donnan membrane theory is discussed in Chapter 14. According to the theory, the ratio of the intracellular/extracellular concentration of a permeant univalent anion is equal to the ratio of extracellular/intracellular concentration of a permeant univalent cation. A more general mathematical expression that includes ions of any valence is

$$\left(\frac{A_i}{A_e}\right)^{1/Z_a} = \left(\frac{C_e}{C_i}\right)^{1/Z_c} = r \qquad (8)$$

where A_i is the intracellular and A_e the extracellular concentration of anion, Z_c is the valence of cation, Z_a is the valence of anion, C_i is the intracellular and C_e the extracellular concentration of cation and r is the Donnan factor. The value of r depends upon the average molecular weight and valence of the macromolecules (mostly protein) within the cell and the intracellular and extracellular volumes. Since the macromolecules within the cell are charged negatively, the cation concentration will be higher within the cell, that is, $C_i > C_e$. Since a Donnan distribution results in a membrane potential, the distribution of drug ion also will be in keeping with the membrane potential.

The Donnan distribution also applies to the distribution of a charged drug between the plasma and interstitial compartment, because of the presence of anionic proteins in the plasma. Eq 8 applies by changing the subscript i to p, for plasma, and e to i, for interstitial. The Donnan factor, r, for plasma–interstitial space partition is about 1.05:1.

Binding and Storage—Drugs frequently are bound to plasma proteins (especially albumin), interstitial substances, intracellular constituents and bone and cartilage. If binding is extensive and firm, it will have a considerable impact upon the distribution, excretion and sojourn of the drug in the body. Obviously, a drug that is bound to a protein or any other macromolecule will not pass through the membrane in the bound form; only the unbound form can negotiate among the various compartments.

The partition among compartments is determined by the binding capacity and binding constant in each compartment. As long as the binding capacity exceeds the quantity of drug in the compartment, the following equation generally applies:

* The relationship of ionization and partition to pH and pK has been formulated in several different ways, but the student may calculate the concentrations from simple mass law equations. More sophisticated calculations and reviews of this subject are available.[6,11–16]

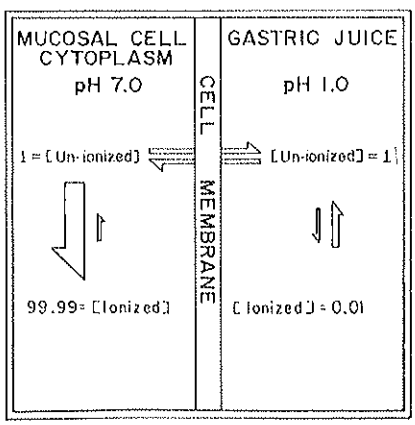


Fig 35-11. Hypothetical partition of salicylic acid between gastric juice and the cytoplasm of a gastric mucosal cell. It is assumed that the ionized form cannot pass through the cell membrane. The intragastric concentration of salicylic acid is arranged arbitrarily to provide unit concentration of the un-ionized form. *Bracketed values:* concentration; *arrows:* relative size depicts the direction in which dissociation-association is favored at equilibrium.

$$\log D_b = \log K + a \log D_f \quad (9)$$

where D_b is the concentration of bound drug, D_f is the concentration of free drug and a and K are constants characteristic of the drug and binding macromolecule. The equation is that of a Freundlich isotherm. As the binding capacity is approached, the relationship no longer holds. For a nondissociable drug at equilibrium, D_f will be the same in all communicating compartments, so that it would be possible to calculate the partition if K and a are known for each compartment. Except for plasma, the values of K and a are generally unknown, hut the percent bound is often known. From the percent bound, the partition also can be calculated, as in Fig 35-12. However, the logarithmic relationships shown in Eq 9 serve as a reminder that the percent bound changes with the concentration, so that the partition will vary with the dose. If the drug is a weak acid or base, the unionized free form negotiates among the compartments, but the ionized form is often the more firmly bound, and calculations must take into account the dissociation constant and the different K's and a's of the ionized and un-ionized forms.

It is misbelieved commonly that binding in the plasma interferes with the activity of a drug and the intracellular binding in a responsive cell increases activity or toxicity. Both binding in plasma and in the tissues decreases the concentration of free drug, but this is easily remedied by adjusting the dose to give a sufficient concentration for pharmacological activity. The distribution and activity of the free form is not affected by binding. The principal effect of binding is to increase the initial dose requirement for the drug and create a reservoir of drug from which the drug may be withdrawn as the free form is excreted or metabolized. However, if the binding is extremely firm and release is slow, the rate of release may not be enough to sustain the free form at a sufficient level for pharmacological activity; in such instances the bound drug cannot be considered a reserve.

The effect of binding upon the sojourn of a drug may be considerable. For example, quinacrine, which may he concentrated in the liver to as much as several thousand times the concentration in plasma, may remain in the body for months. Some iodine-containing radiopaque diagnostic agents are bound strongly to plasma protein and may remain in the plasma for as long as 2 yr. In pathological conditions, such as nephrosis, diahetes or cirrhosis, in which plasma protein levels may be decreased, the plasma protein hinding, loading dose and duration of action all may be decreased.

If a drug is hound to a functional macromolecule, binding may relate to pharmacological activity and toxicity, providing that the binding is at a critical center of the macromolecule. The binding by nucleic acids of certain antimalarials, such as quinacrine, undoubtedly contributes to the parasiticidal actions as well as to toxicity.

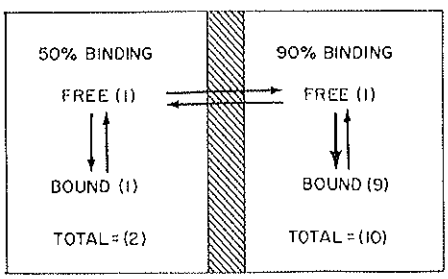


Fig 35-12. Distribution of a drug between two compartments in which the degrees of binding to protein differ. The percent of binding is indicated. Only the unbound drug can pass through the membrane. *Bracketed values:* concentration (courtesy, Schanker[12]).

Most drugs are bound to proteins by relatively weak forces, such as van der Waals (London, Keesom or Debye) forces, or hydrogen or ionic bonds. Consequently, binding constants generally are small and binding is usually readily reversible. The larger the molecule, the greater the van der Waals bonding, so that large drug molecules are more likely to be bound strongly than are small ones.

Just as shape and the nature of functional groups is important to drug–receptor combination, so they also are to binding. Drugs of similar shape and/or chemical affinities may bind at the same sites on a binding protein and hence compete with one another. For example, phenylbutazone displaces warfarin from human plasma albumin, which may cause an increase in the anticoagulant effect of warfarin. Some drugs also may displace protein-bound endogenous constituents. For example, sulfisoxazole displaces bilirubin from plasma proteins; in infants with kernicterus the freed bilirubin floods the central nervous system and causes sometimes fatal toxicity.

Depending on the lipid–water coefficient, a drug may be taken up into fatty tissue. The ratio of concentration in fat, to that in the plasma, will not be the same as dictated by the partition coefficient because of the content of water and nonlipids in adipose tissue, and because electrolytes and other solutes alter the dielectric constant and hence solubilities from those of pure water. Lipoproteins, and even nonpolar substituents on plasma proteins, also take up lipid-soluble molecules, so that solubility in plasma can be considerably higher than that in water. The relatively high solubility of ether in plasma makes plasma a pool for ether, the filling of which delays the onset of anesthesia. However, ether and other volatile anesthetics are taken up gradually into the adipose tissue, which acts as a store of the anesthetic. The longer the anesthetic is administered, the greater the store and the longer it takes for anesthesia to terminate when inhalation has been discontinued.

Another notable substance that is taken up readily into fat is thiopental. Even though there is a high solubility of this barbiturate in fat, the low rate of blood flow in fat limits the rate of uptake. Because the blood flow in the brain is very high, thiopental rapidly enters brain tissue. However, it soon equilibrates with the other tissues, and the brain concentration falls as that in the other tissues (eg, muscle or liver) increases. Gradually, however, the fat accumulates the drug at the expense of other compartments. The gradual entry of thiopental into fat at the expense of plasma, muscle or liver is illustrated in Fig 35-13.

Nonequilibrium and Redistribution—Thus far, the distribution of drugs has been discussed mainly as though equilibrium or steady state conditions exist after a drug is absorbed and distributed. However, since most drugs are administered at intervals and the body content of drug rises and falls with absorption and destruction-excretion, neither

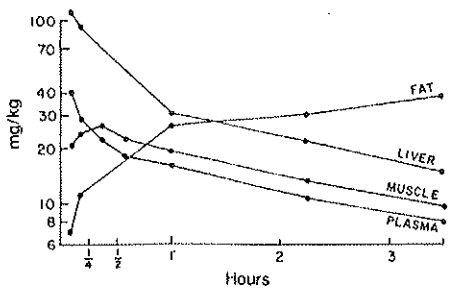


Fig 35-13. Predisposition of thiopental for fat. 25 mg/kg was given to a dog. After a brief sojourn in the more vascular tissues, thiopental gradually transfers to fat, where the lipid-soluble drug dissolves in fat droplets (courtesy, Brodie and Hogben[13]).

a true equilibrium among the body compartments nor a steady state exists.

The term equilibrium is used misleadingly to describe the conditions that exist when the plasma concentration and the concentration in a tissue are equal, as exemplified at the point of intersection of the curves for plasma and muscle or plasma and fat in Fig 35-13. But such "equilibrium" with fat occurs much later than "equilibrium" with muscle, so that no true equilibrium really exists among all the compartments. Furthermore, the crossover point for plasma and any one tissue is not necessarily an equilibrium point, because the rates of ingress and egress from the tissue are not necessarily equal when the internal and external concentrations are equal, since there are numerous factors that make for unequal distribution (pH partition, Donnan effect, electrochemical distribution, active transport, binding, etc).

A study of Fig 35-13 shows that the distribution of thiopental continually changed during the 3.5 hr of observation. At the end of the period, the content in fat was still increasing while that in each of the other compartments was decreasing. This time-dependent shift in partition is called *redistribution*. Eventually, the content in fat would have reached a peak, which would represent as nearly a true equilibrium point as could be achieved in the dynamic situation where metabolic destruction and a slight amount of excretion of the drug was taking place. Once the concentration in the fat had reached its peak, its content would have declined in parallel with that in the other tissues and the partition among the compartments would have remained essentially constant. Redistribution, then, takes place only until the concentration in the slowest filling compartment reaches its peak, so long as the kinetics of elimination are constant.

An index of distribution known as the *volume of distribution* (amount of drug in the body divided by plasma concentration) is of considerable usefulness in pharmacokinetics but is of limited value in defining the way in which a drug is partitioned in the body. Volume of distribution is discussed on page 727.

The word *space* is often used synonymously with volume of distribution. It is employed especially when the distributed substance has a volume of distribution that is essentially identical to a physical real space or body compartment. *N*-acetyl-4-aminoantipyrine is distributed evenly throughout the total body water and is not bound to proteins or other tissue constituents. Thus, the acetylaminoantipyrine space, or volume of distribution, coincides with that of total body water. Inulin, sucrose, sulfate and a number of other substances essentially are confined to extracellular water, so that an inulin space, for example, measures the extracellular fluid volume. Evans blue is confined to the plasma, so that the Evans blue space is the plasma volume. Such space measurements with standard space indicators are a necessary part of studies on the distribution of drugs, since it is desirable to compare the volume of distribution to a drug to the physiological spaces.

Biotransformations

Most drugs are acted upon by enzymes in the body and converted to metabolic derivatives called metabolites. The process of conversion is called *biotransformation*. Metabolites are usually more polar and less lipid-soluble than the parent drug because of the introduction of oxygen into the molecule, hydrolysis to yield more highly polar groups or conjugation with a highly polar substance. As a consequence, metabolites often show less penetrance into tissues and less renal tubular resorption than the parent drug, in accordance with the principle of the low penetrance of polar and high penetrance of lipid-soluble substances. For similar reasons, metabolites are usually less active than the parent drug, often inactive; even if they are appreciably active, they generally are excreted more rapidly. Therefore, the usual net effect of biotransformation may be said to be one of *inactivation* or *detoxication*.

There are, however, numerous examples in which biotransformation does not result in inactivation. Table III (page 742) lists a number of drugs that generate active metabolites; in a few instances activity derives entirely from the metabolite.

There are also examples in which the parent drug has no activity of its own but is converted to an active metabolite: parathion, malathion and certain other anticholinesterases require metabolic activation; inactive chloroguanide is converted to an active triazine derivative; phenylbutazone is hydroxylated to the antirheumatic hydroxyphenylbutazone; inactive pentavalent arsenicals are reduced to their active trivalent metabolites and there are other examples of an activating biotransformation.

When a delayed or prolonged response to a drug is desired or an unpleasant taste or local reaction is to be avoided, it is a common pharmaceutical practice to prepare an inactive or nonoffending precursor, such that the active form may be generated in the body. This practice has been termed *drug latentiation*. Chloramphenicol palmitate, dichloralphenazone and the estolates of various steroid hormones are examples of deliberately latentiated drugs. Because inactive metabolites do not always result from biotransformation, the term detoxication should not be used as a synonym for biotransformation. See Chapter 25.

Biotransformations take place principally in the liver, although the kidney, skeletal muscle, intestine or even plasma may be important sites of the enzymatic attack of some drugs. Since plasma lacks the enzymes and structures required for electron transport, biotransformations in plasma are mostly hydrolytic.

Endoplasmic Reticulum and Microsomal System—Biotransformations in the liver mainly occur in *smooth endoplasmic reticulum*. The endoplasmic reticulum is a tubular system which courses through the interior of the cell but also appears to communicate with the interstitial space, and its membrane is continuous with the cell membrane. Some of the reticulum is lined with ribonucleoprotein particles, called ribosomes, which are engaged in protein synthesis; this is the *rough* endoplasmic reticulum. Although the smooth endoplasmic reticulum lacks such a granular appearance, it is invested heavily with numerous enzymes which biotransform many drugs and some endogenous substances.

When a broken-cell homogenate of the liver is prepared, the reticulum becomes fragmented and the fragments form vesicular structures called *microsomes*. Although the microsomes are artifacts, it is the practice to refer to the *microsomal drug metabolizing system* rather than to the smooth endoplasmic reticulum.

The microsomal system is peculiar in that both oxidations and reductions usually require the reducing cofactor, reduced nicotinamide adenine dinucleotide phosphate (NADPH). This is because microsomal oxidations proceed by way of the introduction of oxygen rather than by dehydrogenation and NADPH is essential to reduce one of the atoms of oxygen. The drug first binds to an oxidized cytochrome P-450. The drug–cytochrome complex then is reduced by NADPH-cytochrome reductase; the reduced complex then combines with oxygen, after which the metabolite is released and oxidized cytochrome P-450 is regenerated. Cytochrome P-450 is a generic term that includes at least 30 and perhaps as many as 100 separate enzymes.[17]

Some of the enzymes of the microsomal system are quite easily *induced;* that is, a substrate of the enzyme may increase considerably the activity of that enzyme by increasing

the biosynthesis of that enzyme. An increase in the amount of smooth endoplasmic reticulum sometimes also occurs concomitantly with enzyme induction.

Treatment of an experimental subject with phenobarbital greatly will increase the rate of metabolism of phenobarbital, which necessitates larger and more frequent doses of the drug in order to maintain a constant sedative effect. Moreover, phenobarbital may induce an increased metabolism of some other, but not all, barbiturates as well as some unrelated drugs, such as strychnine and warfarin. Oddly, warfarin does not induce its own biotransformation readily. At the present time, both self-induction and cross-induction appear capricious and unpredictable.

Induction may create therapeutic problems. For example, the use of phenobarbital during treatment with warfarin increases the dose requirement for warfarin. If the physician is unaware of this interaction and fails to increase the dose, the patient may suffer a thrombotic episode. If the dose of warfarin has been increased and the phenobarbital is then discontinued, the rate of metabolism of warfarin may drop to its previous level, so that the patient is overdosed, with hemorrhagic consequences. Some drugs inhibit rather than induce the microsomal system, which reduces the dose requirement and may lead to toxicity. Cimetidine is an example of a drug that inhibits the hepatic metabolism of a number of other drugs.

The activity of the microsomal system is affected by many factors other than the presence of drugs. Age, sex, nutritional states, pathological conditions, body temperature and genetic factors are among the influences that have been identified. Age, particularly, has received considerable attention. Infants have a poorly developed microsomal system, which accounts for the low dose requirement for morphine and also explains the high toxicity of chloramphenicol in infants.

The activity and selectivity of the microsomal system varies greatly from species to species, so that care must be exercised in extrapolating experimental findings in laboratory animals to man.

Types of Biotransformations—Biotransformations may be *degradative*, wherein the drug molecule is diminished to a smaller structure, or *synthetic*, wherein one or more atoms or groups may be added to the molecule. Very few drugs are degraded completely. However, it is more useful to categorize biotransformations with respect to "metabolic" (nonconjugative) biotransformations and conjugative biotransformations. The former is called phase I and the latter phase II. In phase I, pharmacodynamic activity may be lost; however, active and chemically reactive intermediates also may be generated. The polarity of the molecule may or may not be increased sufficiently to increase excretion markedly. In phase II, metabolites from phase I may be conjugated and sometimes the original drug may be conjugated, thus bypassing phase I. Phase II generates metabolites of high polarity which are excreted readily.

Biotransformations may be placed into five categories: (1) oxidation, (2) reduction, (3) hydrolysis, (4) conjugation and (5) miscellaneous. Oxidation, reduction and hydrolysis comprise phase I. Conjugation comprises phase II. Most of the miscellaneous biotransformations belong in phase I.

Oxidation—Oxidation is more common than any other type of biotransformation. Oxidations that occur primarily in the liver microsomal system include side-chain hydroxylation, aromatic hydroxylation, deamination (which is oxidative and results in the intermediate formation of RCHO), *N*-, *O*-, and *S*-dealkylation (which probably involves hydroxylation of the alkyl group followed by oxidation to the aldehyde) and sulfoxide formation. *N*-Demethylation involves a different system from *N*-dealkylation of higher radicals.

Oxidations that occur elsewhere, other than the microsomes, are generally dehydrogenations followed by the addition of oxygen or water. Examples are the oxidation of alcohols by alcohol dehydrogenase, the oxidation of aldehyde by aldehyde dehydrogenase and the deamination of monoamines by monoamine oxidase and diamines by diamine oxidase. The oxidation of purines, like caffeine and theophylline, is also extramicrosomal.

Reduction—Reductions are relatively uncommon. They mainly occur in liver microsomes, but they occasionally take place in other tissues. Examples are the reduction of nitro and nitroso groups (as in chloramphenicol, nitroglycerin and organic nitrites), of the azo group (as in prontosil) and of certain aldehydes to the corresponding alcohols (as with the deaminated serotonin metabolite, 5-hydroxytryptophal, to 5-hydroxytryptophol).

Hydrolysis—Hydrolysis is a common biotransformation among esters and amides. Esterases are located in many structures besides the microsomes. For example, cholinesterases are found in plasma, erythrocytes, liver, nerve terminals, junctional interstices and postjunctional structures, and procaine esterases are found in plasma. Various phosphatases and sulfatases also are distributed widely in tissues and plasma, although few drugs are appropriate substrates. The hydrolytic deamidation of meperidine occurs primarily in the hepatic microsomes.

Conjugation—A large number of drugs, or their metabolites, are conjugated. Conjugation is the biosynthetic process of combining a chemical compound with a highly polar and water-soluble natural substance to yield a water-soluble, usually inactive, product. Conjugations generally involve either esterification, amidation, mixed anhydride formation, hemiacetal formation or etherization.

Glucuronic acid is the most frequent partner to the drug in conjugation. Actually, the drug reacts with uridine diphosphoglucuronic acid rather than with simple glucuronic acid. The drug or drug metabolite combines at the number 1 carbon (aldehyde end) and not at the carboxyl end of glucuronic acid. The hydroxyl group of an alcohol or a phenol attacks the number 1 carbon of the pyran ring to replace uridine diphosphate. The product is a hemiacetal-like derivative. Since the product is not an ester, the term *glucuronide* is appropriate. Rarely, thiols and amines may form analogous glucuronides.

Carboxyl compounds form esters, appropriately called *glucuronates*, in replacing the uridine diphosphate. *Sulfuric acid* is also a frequent conjugant, especially with phenols and to a lesser extent with simple alcohols. The sulfurated product is called an *ethereal sulfate*. Occasionally sulfuric acid conjugates with aromatic amines to form *sulfamates*. *Phosphoric acid* also conjugates with phenols and aromatic amines. The conjugation of benzoic acid with glycine to yield hippuric acid is a classical example of an *amidation* conjugative process. Cysteine may take the place of glycine, through the intermediatiun of glutathione, to yield mercapturic acids with certain aromatic acids.

Amidations with amino acids are less frequent than *acetylation*, partly because few drugs are carboxylic compounds. Aromatic amines and occasionally aliphatic amines or heterocyclic nitrogen frequently are acetylated. Acetyl-CoA is the biological reagent rather than acetic acid itself. Unlike most other conjugates, the acetylate (amide) is usually less water-soluble than the parent compound. The acetylation of the para-amino group of the sulfonamides is a prime example of this type of conjugation.

Although most conjugations occur in the liver, the microsomal system is not involved. Some conjugations occur in the kidney or in other tissues.

Miscellaneous—Many amines, especially derivatives of β-phenylethylamine and heterocyclic compounds, are methylated in the body. The products are usually biologically active, sometimes more so than the parent compound. *N-Methylation* may occur in the cell sap of the liver and elsewhere, especially in chromaffin tissue in the case of phenylethylamines.

Phenolic compounds may be *O*-methylated. *O-Methylation* is the principal route of biotransformation of catecholamines such as epinephrine and norepinephrine, the methyl group being introduced on the *meta*-hydroxy substituent. Both *N*- and *O*-methylation require *S*-methyladenosyl cysteine.

Desulfuration, in which oxygen may replace sulfur, takes place in the liver. Thiopental is converted in part to pentobarbital by desulfuration, and parathion is transformed to paraoxon.

Dehalogenation of certain insecticides and various halogenated hydrocarbons may take place, principally in the liver but not in the microsomes.

Excretion

Some drugs are not biotransformed in the body. Others may be biotransformed, but their products still remain to be eliminated. It follows that excretion is involved in the elimination of all drugs and/or their metabolites. Although the kidney is the most important organ of excretion, some substances are excreted in bile, sweat, saliva, gastric juice or from the lungs.

Renal Excretion—The excretory unit of the kidney is called the *nephron* (Fig 35-14). There are several million nephrons in the human kidney. The nephron is essentially

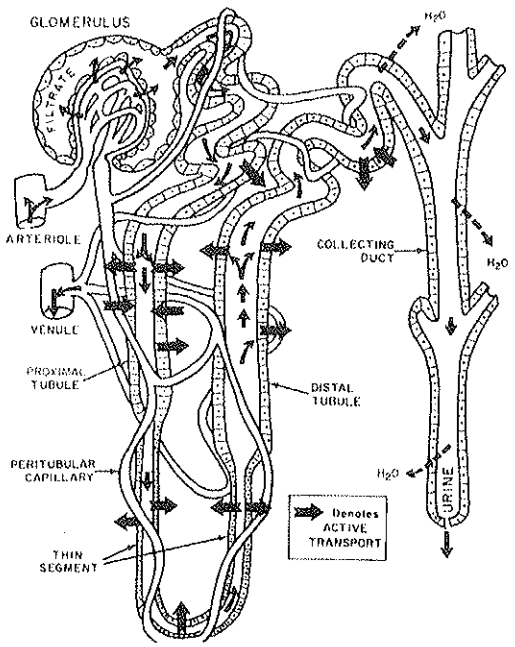


Fig 35-14. Diagram of a mammalian nephron. Note how the lower loops of the postglomerular capillaries course downward and double back along with the tubule. This allows countercurrent distribution to maintain hyperosmolar urine within the thin segment.

a filter funnel, called *Bowman's capsule*, with a long stem, called a *renal tubule*. It also is recognized now that the collecting duct is functionally a part of the nephron. The *blood vessels* that invest the capsule and the tubule are also an essential part of the nephron.

Bowman's capsule is packed with a tuft of branching interconnected capillaries (*glomerular tuft*), which provide a large surface area of capillary endothelium ("filter paper") through which fluid and small molecules may filter into the capsule and begin passage down the tubule. The glomerular tuft, together with Bowman's capsule, constitute the *glomerulus*. The glomerular capillary endothelium and the supporting layer of Bowman's capsule have channels ranging upward to 40 Å. Consequently, all unbound crystalloid solutes in plasma, and even a little albumin, pass into the glomerular filtrate.

The postglomerular vessels, which lie close to the tubules, are critically important to renal function in that substances resorbed from the filtrate by the tubule are returned to the blood along these vessels. The tubule is not straight but rather first makes a number of convolutions (called a *proximal convoluted tubule*), then courses down and back up a long loop (called the *loop of Henle*), makes more convolutions (the *distal convoluted tubule*) and finally joins the collecting duct. The loop of Henle is divided into a *proximal (descending) tubule*, a thin segment and a *distal (ascending) tubule*.

As the glomerular filtrate passes through the proximal tubule, some solute may be resorbed (*tubular resorption*) through the tubular epithelium and returned to the blood. Resorption occurs in part by passive diffusion and in part by active transport, especially with sodium and glucose. Chloride follows sodium obligatorily.

In the proximal region, the tubule is quite permeable to water, so that resorbed solutes are accompanied by enough water to keep the resorbate isotonic. Consequently, although the filtrate becomes diminished in volume by approximately 80% in the proximal tubule, it is not concentrated.

Some *acidification* occurs in the proximal tubule as the result of carbonic anhydrase activity in the tubule cells and the diffusion of hydronium ions into the lumen. In the lumen the hydronium ion reacts with bicarbonate ion, which is converted to resorbable nonionic CO_2.

There is also active transport of organic cations and anions into the lumen (*tubular secretion*), each by a separate system. These active transport systems are extremely important in the excretion of a number of drugs; for example, penicillin G is secreted rapidly by the anion transport system and tetraethylammonium ion by the cation transport system. Probenecid is an inhibitor of anion secretion and, hence, decreases the rate of loss of penicillin from the body.

As the filtrate travels through the thin segment it becomes concentrated, especially at the bottom, as a result of active resorption and a countercurrent distribution effect enabled by the recurrent and parallel arrangement of the ascending segment, the parallel orientation of the collecting duct and the similar recurrent geometry of the associated capillaries.

In the thick segment of the ascending loop of Henle, both sodium and chloride are transported actively.

In the distal tubule, sodium resorption occurs partly in *exchange* for potassium (*potassium secretion*) and for hydronium ions. Adrenal mineralocorticoids promote distal tubular sodium resorption and potassium and hydronium secretion. *Ammonia secretion* also occurs, so that the urine either may be acidified or alkalinized, according to acid–base and electrolyte requirements.

Water is resorbed selectively from the distal end of the distal convoluted tubule and the collecting ducts; water resorption is under the control of the antidiuretic hormone.

Drugs also may be resorbed in the distal tubule; the pH of the urine there is extremely important in determining the rate of resorption, in accordance with the principle of nonionic diffusion and pH partition. The pH of the tubular fluid also affects the tubular secretion of drugs.

As an example of the importance of urine pH, in humans the secondary amine, mecamylamine, is excreted more than four times faster when the urine pH is less than 5.5 than when it is above 7.5; Fig 35-15 illustrates the effect of urine pH on the excretion of this amine. The effect of urine pH on the excretion of a weak acid, sulfaethidole (for the structure, see page 1109, RPS-15), is shown in Fig 35-16.

The urine pH and, hence, drug excretion may fluctuate widely according to the diet, exercise, drugs, time of day and other factors. Obviously, the excretion of weak acids and bases can be controlled partly with acidifying or alkalinizing salts, such as ammonium chloride or sodium bicarbonate, respectively. Comparative studies on potency and efficacy

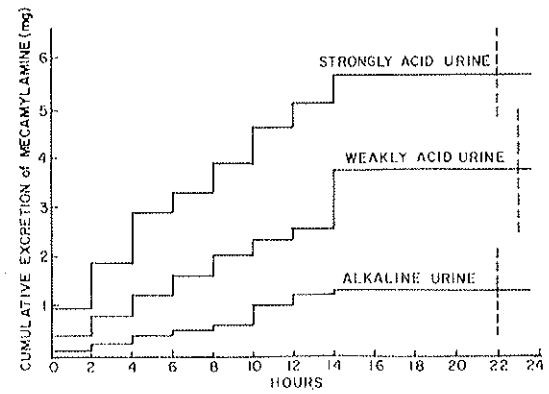


Fig 35-15. The effect of urinary pH on the mean cumulative excretion in man of mecamylamine during the first day after oral administration of 10 mg. *Vertical broken lines:* standard deviation (courtesy, Milne, *et al*[18]).

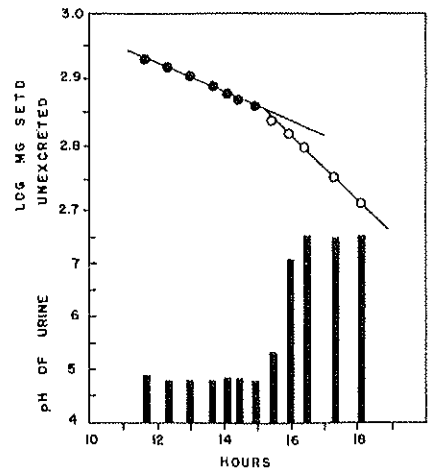


Fig 35-16. The effect of urinary pH on the excretion of sulfaethidole in a human subject after oral administration of 2 g. *Bars* (lower half): urinary pH; *circles* (open and closed, top): log of the amount of drug remaining in the body; *negative slopes* (of lines defined by the circles): a function of the rate constant of excretion. Note the abrupt increase in rate when the urinary pH is changed from acidic to neutral or slightly alkaline (courtesy, Kostenbauder, *et al*[19]).

in man have demonstrated the importance of controlling urinary pH. Urine pH is important only when the drug in question is a weak acid or base of which a significant fraction is excreted. The plasma levels will change inversely to the excretory rate. For example, it has been shown clinically with quinidine that alkalinization of the urine not only decreases the urine concentration but also increases the plasma concentration and toxicity.

The collecting duct also resorbs sodium and water, secretes potassium, acidifies and concentrates the urine. Antidiuretic hormone (ADH) controls the permeability to water of both the collecting duct and the distal tubule.

Renal clearance and the kinetics of renal elimination are discussed in Chapter 36 (page 730).

Biliary Excretion and Fecal Elimination—Many drugs are secreted into the bile and, thence, pass into the intestine. A drug that is passed into the intestine via the bile may be reabsorbed and not lost from the body. This cycle of biliary secretion and intestinal resorption is called *enterohepatic circulation*. Examples of drugs enterohepatically circulated are morphine and the penicillins. The biliary secretory systems greatly resemble those of the kidney tubules. The enterohepatic system may provide a considerable reservoir for a drug.

If a drug is not absorbed completely from the intestine, the unabsorbed fraction will be eliminated in the feces. An unabsorbable drug that is secreted into the bile will likewise be eliminated in the feces. Such fecal elimination is called *fecal excretion*. Only rarely are drugs secreted into the intestine through the succus entericus (intestinal secretions), although a number of amines are secreted into gastric juice.

Alveolar Excretion—The large alveolar area and high blood flow make the lungs ideal for the excretion of appropriate substances. Only volatile liquids or gases are eliminated from the lungs. Gaseous and volatile anesthetics essentially are eliminated completely by this route. Only a small amount of ethanol is eliminated by the lungs, but the concentration in the alveolar air is related so constantly to the blood alcohol concentration that the analysis of expired air is acceptable for legal purposes. The high aqueous solubility and relatively low vapor pressure of ethanol at body temperature account for the retention of most of the substance in the blood. Carbon dioxide from those drugs that are partly degraded also is excreted in the lungs.

Pharmacokinetics

Pharmacokinetics is the science that treats of the rate of absorption, extent of absorption, rates of distribution among body compartments, rate of elimination and related phenomena. Because of its importance, two chapters, *Basic Pharmacokinetics* (page 725) and *Clinical Pharmacokinetics* (page 746), have been devoted to the subject.

Drug Interaction and Combination

Frequently a patient may receive more than one drug concurrently. Case records show that surgical patients commonly receive more than ten, and sometimes as many as 30, drugs and the patient is often under the influence of several drugs at once, sometimes unnecessarily. Multiple-drug administration also is common for patients hospitalized for infections and other disorders. Furthermore, a patient may be suffering from more than one unrelated disorder which demands simultaneous treatment with two or more drugs. In such instances, interactions are unsolicited and often unexpected.

In addition to the administration of drugs concurrently for their independent and unrelated effects, drugs are sometimes administered concurrently deliberately to make use of expected interactions.

Types of Interaction and Reasons for Combination Therapy

A drug may affect the response to another drug in a quantitative way. On the one hand, the intensity of either the therapeutic effect, or side effect, may be augmented or suppressed. On the other hand, a qualitatively different effect may be brought out. The mechanisms of such interactions are many and not always are understood. A drug may not necessarily affect either the quality or initial intensity or effect of another drug, but may cause significant to profound changes in the duration of action. The nature of this type of interaction generally is understood fairly well, although it may not yet have been ascertained for any particular drug combination. The deliberate use of combined interacting drugs is most valid when the mechanism of the interaction is understood and the combined effects are both quantifiable and predictable. The rationales of drug combination and the principles involved are discussed below.

Combinations to Increase Intensity of Response or Efficacy—Sometimes the basis for the action of one drug to increase the intensity of response to another is well understood, but often the reason for a positive interaction is obscure. A terminology has arisen that frequently is not only enlightening as to mechanisms and principles but also which is somewhat confusing.

Drugs that elicit the same quality of effect and are mutually interactive are called *homergic*, regardless of whether there is anything in common between the separate response systems. Thus, the looseness of the term admits a pressor response consequent to an increase in cardiac output to be

homergic with one resulting from arteriolar constriction, even though there is not one common responsive element, the blood pressure itself being but a passive indicator. However, homergic drugs usually have in common at least part of a response system. Thus, both norepinephrine and pitressin stimulate some of the same vascular smooth muscle, even though they do not excite the same receptors.

Two homergic drugs can be agonists of the same receptor, so that the entire response system is common to both. Such drugs are called *homodynamic*. As discussed under *Drug Receptors and Receptor Theory* (page 702), homodynamic drugs will generate dose–intensity of effect curves with parallel slopes, but not necessarily with identical maxima or efficacies, if one of the drugs is a partial agonist.

From mass-law kinetics and dose–effect data of the separate drugs, it is possible to predict the combined effects of two agonists to the same receptor. If both drugs are full agonists, theory predicts that an *EDx* of Drug *A* added to an *EDy* of Drug *B* should elicit the same effect as that of an *EDy* of Drug *A* added to an *EDx* of Drug *B*. An example is shown in Fig 35-17. Dose–percent data with homodynamic drugs can be treated in the same way.[21]

Drugs whose combined effects fit the above conditions are called *additive*. If the response to the combination exceeds the expected value for additivity, the drugs are considered to be *supra-additive*. Purely homodynamic drugs do not show supra-additivity; however, if one drug in the pair has an additional action to affect the concentration or penetrance of the other or to prime the response system in some way, two agonists to the same receptor may exhibit supra-additivity. Two homergic drugs are *infra-additive* if their combined effect is less than expected from additivity. As with supra-additivity, infra-additivity must involve an action elsewhere than on a common receptor.

Two drugs are said to be *summative* if a dose of drug that elicits response x added to a dose of another drug that elicits response y gives the combined response $x + y$. Very little significance usually can be attached to summation. Unless the dose–intensity curve of each drug is linear, rather than log-linear, summation cannot be predicted from the two curves. When summation does occur with the usual clinical doses of two drugs, it almost never occurs over the entire dose range; indeed, if the dose of each of the two drugs is greater than an ED50, summation is theoretically impossible unless it is possible to increase the maximal response. At best, summation is an infrequent clinical finding limited to one or two doses.

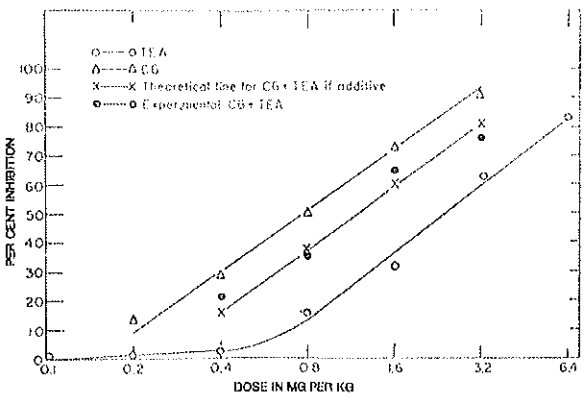


Fig 35-17. Additive inhibitory effects of totraethylammonium (TEA) and hexamethonium (C6) on the superior cervical ganglion of the cat. The theoretical line for additivity was calculated on the basis that an increment of TEA added to an EDx of C6 should have the same effect as if it were added to an EDx of TEA. When TEA and C6 were administered together, an equal amount of each was given. The dose is the sum of the doses of the two components (courtesy, Harvey[20]).

Two drugs are said to be *heterergic* if the drugs do not cause responses of the same quality. When heterergy is positive, ie, the response to one drug is enhanced by the other, *synergism* is said to occur. The word often has been used to describe any positive interaction, but it should be used only to describe a positive interaction between heterergic drugs. The term *potentiation* has been used synonymously with synergism, but misuse of the term has led to the recommendation that the term be dropped. Synergism is often the result of an effect to interfere with the elimination of a drug and, thus, to increase the concentration; synergism also may result from an effect on penetrance or on the responsivity of the effector system. Examples of a synergistic effect, in which responsivity is enhanced, are the action of adrenalcorticoids to enhance the vasoconstrictor response to epinephrine and the increase of epinephrine-induced hyperglycemia consequent to impairment by theophylline of the enzymatic destruction of the cAMP which mediates the response.

In clinical practice two homodynamic drugs rarely are coadministered for the purpose of increasing the response, since a sufficient dose of either drug should be able to achieve the same effect as a combination of the two. Most clinical combinations with positively interacting drugs are with heterergic drugs.

Combinations to Decrease Individual Doses and Toxicity—When homodynamic drugs are coadministered, it is usually for the purpose of decreasing toxicity. If the toxicities of two homodynamic drug are infra-additive, the toxicity of combined partial doses of the two drugs often will be less than with full doses of either drug. This principle is valid for trisulfapyrimidines mixture (see page 1181).

Combinations to Attack a Disease Complex at Different Points—With many diseases, more than one organ or tissue may be affected or events at more than one locus may bear upon the ultimate perturbation. For example, in duodenal ulcer, psychic factors appear to increase activity in the vagus nerve, which modulates gastric secretion, so that it is rational to explore the effects of sedatives, ganglionic blocking drugs, antimuscarinic drugs and antacids, singly and in combination. In heart failure the decrement in renal plasma flow and changes in aldosterone levels promote the retention of salt and water, so that diuretics and digitalis usually are employed concomitantly. Pain, anxiety and agitation or depression are frequent accompaniments of various pathologic processes, so that it is to be expected that analgesics, tranquilizers, sedatives or antidepressives frequently will be given at the same time, along with other drugs intended to correct the specific pathology.

Combinations to Antagonize Untoward Actions—The side effects of a number of drugs can be prevented, or suppressed, by other drugs. An antagonist may compete with the drug at the receptor that initiates the side effect, depress the side-effector system at a point other than the receptor, or stimulate an opposing system.

Antagonism at the receptor is *competitive antagonism* if the antagonist attaches at the same receptor group as the agonist (see page 703). Antagonism at a different receptor group, or inhibition elsewhere in the response system, is *noncompetitive antagonism*. Both competitive and noncompetitive antagonism are classified as *pharmacological antagonism*. The stimulation of an opposing system is *physiological antagonism*.

Examples of pharmacological antagonism are the use of atropine to suppress the muscarinic effects of excess acetylcholine consequent to the use of neostigmine and the use of antihistaminics to prevent the effects of histamine liberated by tubocurarine. Examples of physiological antagonism are

the use of amphetamine to correct partially the sedation caused by anticonvulsant doses of phenobarbital and the administration of ephedrine to correct hypotension resulting from spinal anesthesia.

Combinations That Affect Elimination—Only a few drugs presently are used purposefully to elevate or prolong plasma levels by interfering with elimination, although continued interest in such drugs probably will increase the number.

Probenecid, which already has been mentioned to antagonize the renal secretion of penicillin, was introduced originally for this purpose. However, because penicillin G is inexpensive and available in repository forms, as well as oral forms (obviating the need for injection), it is less imperative to retard the excretion of penicillin. The low nonallergenic toxicity of penicillin permits very large doses to be given without concern for the high plasma concentrations that result, which also means that there is little necessity for increasing the biological half-life of the drug. Consequently, probenecid is not used routinely today in combination with penicillin.

The use of vasoconstrictors to increase the sojourn of local anesthetics at the site of infiltration continues, but few other clinical examples of the deliberate use of one drug to interfere with either the distribution or elimination of another can be cited. Nevertheless, the subject of the effect of one drug on the elimination of another has become immensely active. Innumerable drugs affect the fate of others and the therapist must be aware of such interactions.

Drugs that induce cytochrome P-450s enhance the elimination of drugs that are metabolized by the liver microsomes. There would be very little point ordinarily to solicit combinations that would shorten the duration of action or lower plasma levels, unless it were to reduce an overdosage. However, since such combinations are used unwittingly or unavoidably, this type of interaction is of great clinical importance.

Combinations to Alter Absorption—In the section on *Vehicles and Absorption Adjuvants* (page 714) it was mentioned that certain substances facilitate the absorption of others. The use of such absorption adjuvants generally is included under the subject of formulation rather than under drug combination. Although drugs which increase blood flow, motility, etc have an effect to increase the rate of absorption, the use of such drugs so far has not proved to be very practical. When it is desired to slow the absorption of drugs, various physical or physicochemical means prove to be more effective and less troublesome than drug combinations.

Fixed Combinations of Drugs

Concomitant treatment with two or more drugs frequently is unnecessary, and it, generally, immeasurably complicates therapy and the evaluation of response and toxicity. Nevertheless, it is often warranted, even essential and cannot be condemned categorically. However, with fixed-dose or fixed-ratio combinations, in which the drugs are together in the same preparation, there are certain disadvantages, except for a few rare instances like trisulfapyrimidines.

The disadvantages are as follows: patients differ in their responsivity or sensitivity to drugs and adjustments in dosage or dose-interval may be necessary. If adjustment of only one component of the mixture is required, it is undesirable that the schedule of the second component be adjusted obligatorily, as it is in a fixed combination. According to which way the dose is adjusted, either toxicity or loss of the therapeutic effect may result. Furthermore, when adverse effects to either component occur, both drugs must be discontinued. The fixed combination denies the physician flexible control of therapy. Especially when one component in a mixture is superfluous yet potentially toxic, as is often the case, the promotion of fixed combinations is reprehensible. However, the separate administration of drugs used in combination often complicates treatment for the patient, who, in an outpatient situation and sometimes in the hospital, may not take all of his medication or who may take it at inappropriate intervals. The resulting consequences may be worse than those of fixed combinations in certain instances. Consequently, a summary dismissal of fixed combinations is unwarranted. Rather, the fundamentals of pharmacokinetics and clinical experience must be brought together with biopharmaceutics to analyze present combinations and to predict possible new allowable combinations.

Dangers in Multiple-Drug Therapy

Some objections to fixed-dose combinations were stated above. Also the unanticipated effects of drug combinations have been touched upon, particularly with respect to effects upon elimination. But it should be made clear that more is at stake than simply the biological half-life of a drug. On page 717 an example was given of the grave clinical consequences of the effect of phenobarbital to enhance the biotransformation of warfarin. Other examples of dangerous interactions, such as the effect of several antidepressants greatly to synergize catecholamines, may be cited. Even some antibiotics antagonize each other and increase mortality.

In addition to the obvious pitfalls posed by the interactions themselves, the use of multiple-drug therapy fosters careless diagnosis and a false sense of security in the number of drugs employed. Multiple-drug therapy should never be employed without a convincing indication that each drug is beneficial beyond the possible detriments or without proof that a therapeutically equivocal combination is definitely harmless. Finally, the expense to the patient warrants consideration.

References

1. Clark AJ: *J Physiol (London) 61:* 547, 1926.
2. Ariens EJ, ed: *Molecular Pharmacology*, vol 1, Academic, New York, 176–193, 1964.
3. Stephenson RP: *Brit J Pharmacol 11:* 379, 1956.
4. Rang HP: *Brit J Pharmacol 48:* 475, 1973.
5. Colquhoun D: The relation between classical and cooperative models for drug action. In Rang HP, ed: *Drug Receptors:* University Park, Baltimore, 1973.
6. Schanker LS: *Advan Drug Res 1:* 71, 1964.
7. Brodie BB, *et al: J Pharmacol Exp Ther 130:* 20, 1960.
8. Truitt EB, *et al: J Pharmacol Exp Ther 100:* 309, 1950.
9. Lillehei JP: *JAMA 205:* 531, 1968.
10. Martin AN, *et al: Physical Pharmacy*, 2nd ed, Lea & Febiger, Philadelphia, 247, 253, 1969.
11. Jacobs MH: *Cold Spring Harbor Symp Quant Biol 8:* 30, 1940.
12. Schanker LS: *Pharmacol Rev 14:* 501, 1961.
13. Brodie BB, Hogben CA: *J Pharm Pharmacol 9:* 345, 1957.
14. Hogben CA: *Fed Proc 19:* 864, 1960.
15. Albert A: *Pharmacol Rev 4:* 136, 1952.
16. Ariens EJ, *et al.* In Ariens, EJ, ed. *Molecular Pharmacology*, vol 1, Academic, New York, 7–52, 1964.
17. Whitlock JP Jr: *Ann Rev Pharmacol Toxicol 26:* 333, 1986.
18. Milne MD, *et al: Clin Sci 16:* 599, 1957.
19. Kostenbauder HB, *et al: J Pharm Sci 51:* 1084, 1962.
20. Harvey SC: *Arch Intern Pharmacodyn 114:* 232, 1958.
21. Weaver LC, *et al: J Pharmacol Exp Ther 113:* 359, 1955.

Bibliography

Albert A: *Selective Toxicity*, 7th ed, Chapman and Hall, London, 1985.
Aronow L, *et al: Principles of Drug Action*, 3rd ed, Wiley, New York, 1988.
Barlow RB: *Quantitative Aspects of Chemical Pharmacology*, University Park, Baltimore, 1980.
Benford D, *et al: Drug Metabolism. From Molecules to Man*, Taylor & Francis, New York, 1987.

Black, JW, *et al*, eds: *Perspectives on Receptor Classification*, Liss, New York, 1987.
Boeynaems JM, Dumont JE, eds: *Outlines of Receptor Theory*, Elsevier/North Holland, Amsterdam, 1980.
Boobis AR, *et al*, eds: *Microsomes and Drug Oxidations*, Taylor & Francis, New York, 1985.
Brodie BB: Physicochemical factors in drug absorption. In Binns TB, ed: *Absorption and Distribution of Drugs*, Williams & Wilkins, Baltimore, 1964.
Burgen ASV, Roberts GCK, eds: *Topics in Molecular Pharmacology*, vols 1 & 2, Elsevier, Amsterdam, 1981, 1983.
Colquhoun D: The link between drug binding and response: theories and observations. In O'Brien RD, ed: *The Receptors: A Comprehensive Treatise*, vol 1, Plenum, New York, 1979.
Coulson CJ: *Mechanisms of Drug Actian*, Taylor & Francis, New York, 1987.
DeRobertis E: *Synaptic Receptors, Isolation and Biology*, Dekker, New York, 1975.
Featherstone RM, ed: *A Guide to Molecular Pharmacology*, Parts I and II, Dekker, New York, 1973.
Finean JB, Michell RH, eds: *Membrane Structure*, Elsevier/North Holland, Amsterdam, 1981.
Gilman AG: Guanine nucleotide regulatory proteins and dual control of adenylate cyclase. *J Clin Invest 73*: 93, 1979.
Gregoriadis G, Senior J: *Targeting of Drugs with Synthetic Systems*, Plenum, New York, 1986.
Jakoby WB, *et al*: *Metabolic Basis of Detoxification*, Academic, New York, 1982.
Jenner P, Testa B, eds: *Concepts in Drug Metabolism*, Parts A & B, Dekker, New York, 1980, 1981.
Jolles G, Woolridge KRH, eds: *Drug Design: Fact or Fancy?* Academic, London, 1984.
Kalow W, *et al*, eds: *Ethnic Differences in Reactions to Drugs and Xenobiotics*, Liss, New York, 1986.
Karlin A: Anatomy of a receptor. *Neurosci Comment 1*: 111, 1983.
Kenakin TP: The classification of drugs and drug receptors in isolated tissues. *Pharmacol Rev 36*: 165, 1984.
Kenakin TP: *Pharmacological Analysis of Drug Receptor Interaction*, Raven, New York, 1987.
Lamble JW, ed: *Towards Understanding Receptors*, Elsevier/North Holland, Amsterdam, 1981.
Lamble JW, ed: *More About Receptors*, Elsevier/North Holland, Amsterdam, 1982.
Lamble JW, Abbott AC, eds: *Receptors Again!* Elsevier, Amsterdam, 1984.
Lefkowitz RJ, ed: *Receptor Regulation*, Chapman & Hall, London, 1981.
Lefkowitz RJ, *et al*: Mechanism of hormone-receptor-effector coupling: the β-adrenergic receptor and adenylate cyclase. *Fed Proc 41*: 2664, 1982.
Levine RR: Pharmacology. In *Drug Actions and Reactions*, 3rd ed, Little, Brown, New York, 1983.
Limbird LE: *Cell Surface Receptors: A Short Course on Theory and Methods*, Nijhoff, Boston, 1986.
Makrikanis A: *New Methods in Drug Research*, Prous, Barcelona, 1985.
Martonosi AN: *Membranes and Transport*, Plenum, New York, 1982.
Meyer UA: Role of genetic factors in the rational use of drugs (Chap 18). In Melmon KL, Morrelli HF, eds: *Clinical Pharmacology*, 2nd ed, Macmillan, New York, 1978.
Mitchell JR, Horning MG, eds: *Drug Metabolism and Drug Toxicity*, Raven, New York, 1984.
Nebert DW, *et al*: Genetic mechanisms controlling the induction of polysubstrate monooxygenase (P-450) activities. *Ann Rev Pharmacol Toxicol 21*: 431, 1981.
Olson RW, Venter JC, eds: *Benzodiazepine/GABA Receptors and Chloride Channels*, Liss, New York, 1986.
Ortiz de Montellano PR, ed: *Cytochrome P450. Structure, Mechanism, and Biochemistry*, Plenum, New York, 1986.
Post G, Crooke ST, eds: *Mechanisms of Receptor Regulation*, Plenum, New York, 1986.
Putney JW Jr, ed: *Phosphoinositides and Receptor Mechanisms*, Liss, New York, 1986.
Riethbrock N, Woodcock BG: *Progress in Drug Protein Binding*, Heyden, Philadelphia, 1981.
Roberts, GCK: *Drug Action at the Molecular Level*, University Park, Baltimore, 1977.
Roche EB, ed: *Bioreversible Carriers in Drug Design*, Pergamon, New York, 1987.
Roth SH, Miller KW, eds: *Molecular and Cellular Mechanisms of Anesthetics*, Plenum, New York, 1986.
Sandler M, ed: *Enzyme Inhibitors as Drugs*, University Park, Baltimore, 1980.
Schanker L: Drug absorption. In La Du, *et al*, eds: *Fundamentals of Drug Metabolism and Drug Disposition*, Williams & Wilkins, Baltimore, 1971.
Schmucker DL: Aging and drug disposition. *Pharmacol Rev 37*: 133, 1985.
Schou JS *et al*, eds: *Drug Receptors and Dynamic Processes in Cells*, Raven, New York, 1986.
Stein WD: *Transport and Diffusion Across Cell Membranes*, Academic, Orlando, 1986.
Stenlake JB: *The Chemical Basis of Drug Action*, Athlone, London, 1979.
Strange PG, ed: *Cell Surface Receptors*, Norwood, Chichester, 1983.
Stroud RM: Acetylcholine receptor structure. *Neurosci Comment 1*: 124, 1983.
Testa, B, ed: *Advances in Drug Research*, vols 14, 15, Academic London, 1985, 1986.
Triggle DJ, Janis RA: Calcium channel ligands. *Ann Rev Pharmacol Toxicol 27*: 347, 1987.
Usdin E, *et al*: *Neuroreceptors*, Wiley & Sons, Chichester, 1981.
Van Rossum JM, ed: *Kinetics of Drug Action*, Springer-Verlag, Berlin, 1977.
Venter, JC, Harrison LC, eds: *Molecular and Chemical Characterization of Membrane Receptors*, Liss, New York, 1984.
Vesell E: Pharmacogenetics. *Biochem Pharmacol 24*: 445, 1975.
Vesell ES: Pharmacogenetics—multiple interactions between genes and environment as determinants of drug response. *Am J Med 61*: 83, 1979.
Vesell ES: The influence of host factors on drug response. I. ethnic background. *Ration Drug Ther 13*(8): 1, 1979.
Vesell ES: The influence of host factors on drug response. III. diet. *Ration Drug Ther 14*(5): 1, 1980.
Wardle EN: *Cell Surface Science in Medicine and Pathology*, Elsevier, New York, 1985.

CHAPTER 36

Basic Pharmacokinetics

Stewart C Harvey, PhD
Professor of Pharmacology

and C Dean Withrow, PhD
Associate Professor of Pharmacology
School of Medicine, University of Utah
Salt Lake City, UT 84132

Pharmacokinetics is the discipline which is concerned with the rates of movement of a drug or its metabolites into the body, among its many compartments, out of the body and also which attempts to evaluate the rates of biotransformations of the drug and its metabolites. As in chemistry, it involves primarily following the rate of change in concentration in the appropriate compartment(s), most often in the extracellular fluid (plasma) and/or urine. However, pharmacokinetics is by no means limited to observations of concentration; rates of movement of a drug can be followed by isotopes or other means. The application of pharmacokinetics to drug formulation and treatment regimens also is within the scope of this title. The applications to treatment regimens and other clinical uses of pharmacokinetics are treated in Chapter 37, *Principles of Clinical Pharmacokinetics.*

Orders of Processes

The order of any process is determined by the probability that the appropriate unit events will occur in a given population within a given time. Processes may be zero-order, first-order, second-order, etc, depending upon the number of variables that determine the probability. In pharmacokinetics, only zero-order and first-order processes are important, the latter being of overwhelming significance; consequently, only the kinetics of these two processes will be treated in this chapter.

First-Order Processes

When activity is random within a population of a single species, the probability that a given event will occur is directly proportional to the size of the population. For example, the probability that some atom in a population of radionuclides will disintegrate in any instant is directly proportional to the number of radionuclide atoms in the population. Similarly, the number of molecules of drug that diffuse across a given boundary (eg, the vascular endothelium) per unit time will be proportional directly to the number of molecules near the boundary, which, in turn, is proportional to the concentration. This is the basis of Fick's Law of Diffusion (page 208). Any process in which the rate of change in a population is directly proportional to the population is known as a *first-order* process. In such a process, the time-dependent change in concentration is defined by the equation

$$C = C_0 e^{-kt} \qquad \text{[units of wt} \cdot \text{vol}^{-1} \text{ or molar, etc]} \qquad (1)$$

where C is the concentration at time t, C_0 is the initial concentration (time zero), t is time, e is the natural (Naperian) log base and k is a proportionality constant known as the rate constant. (For a derivation of Eq 1, see page 247.) In a diffusion process, the magnitude of k is determined by the temperature, mobility, permeability and other factors. The numerical value of k also will depend upon the time units (min vs hr, etc) chosen.

Eq 1 predicts that as t approaches infinity, C approaches zero, which would be true for irreversible processes like radioactive decay, diffusion into infinite space, some exentropic SN_1 chemical decompositions and certain enzymatic reactions. However, in a confining space, diffusion and many chemical reactions reach an equilibrium state in which C approaches a finite asymptote as t approaches infinity. Figure 36-1 illustrates a simple situation in which the asymptote is necessarily finite. To satisfy the conditions of this

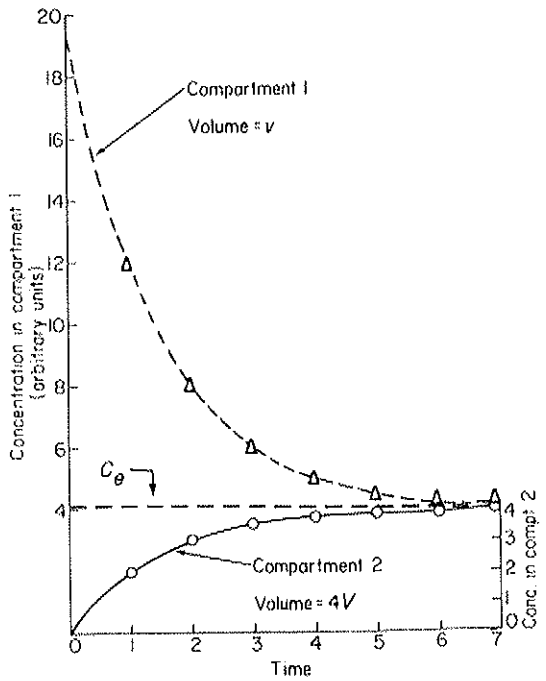


Fig 36-1. Idealized diffusion kinetics of a hypothetical drug that equilibrates between two compartments. Transfer is from compartment 1 into compartment 2. The equilibrium concentration is ⅕ of that initially in compartment 1, because the final volume of distribution is 5 times that of compartment 1.

closed system, $(C_0 - C_e)$ must be substituted for C_0 in Eq 1, C_e being the equilibrium concentration.

In Eq 1, the algebraic sign of k is usually negative, which indicates a diminishing concentration with time. However, in Fig 36-1 the concentration in compartment 2 rises logarithmically with time; nevertheless, k is negative, since the rate diminishes exponentially with time. The equation for the logarithmically rising concentration in compartment 2 will take the form of Eq 5 (page 727), in which C_e would be used in lieu of C_p^∞.

Eq 1 can be written in the log form,

$$\log C = \log C_0 - 0.434kt \qquad \text{[no units]} \qquad (2)$$

The coefficient 0.434 results from the conversion of the natural log base, e, to log base 10 (0.434 = 1/2.303). The equation determines that a plot of log C against t will be rectilinear (bottom of Fig 36-1) with a slope of $-0.434k$ and an ordinate-intercept of C_0. For pharmacokinetics, this is a useful type of plot, because, in the straight-line form, back extrapolation to estimate C_0 is easier and more accurate than from a curve, and k can also be determined graphically.

Rate Constants and Half-Life—Since first-order processes are characterized by exponential or logarithmic kinetics, it follows that a constant fraction of the present or instantaneous population (eg, concentration) changes per unit time, that fraction being equal to $0.434k$; k has the units of t^{-1}. Another way of expressing the rate of change is that of half-time (or especially *half-life*, if the population is decreasing), with the notation $t_{1/2}$. The half-time is the time that it takes the population to decrease (or increase) by 50% of the total possible change. By setting C equal to $\frac{1}{2}C_0$ in either Eqs 1 or 2 and solving for t (which is $t_{1/2}$ under these constraints),

$$t_{1/2} = \frac{0.693}{k} \qquad \text{[units of time]} \qquad (3)$$

Zero-Order Processes

When an enzyme or transport system is saturated, the activity cannot be increased further by increases in the concentration of substrate. Consequently, the rate remains constant so long as the concentration of substrate is in excess of the saturating concentration. In this situation, the rate is independent of the concentration. The kinetics are described as being of *zero-order*, and it is customary to speak of the process as being a zero-order process. The equation describing zero-order kinetics is

$$C = C_0 - kt \qquad [\text{conc vol}^{-1}] \qquad (4)$$

where k has the units of amount/unit time. A plot of C against t on Cartesian coordinates will yield a straight line; a plot of log C against t will yield a curved line. As the process continues, the concentration eventually will fall to subsaturation levels, and the kinetics will change, usually to first-order kinetics, so that it is more appropriate to speak of the initial kinetics and not the process as being zero-order.

Pharmacokinetic Models

The plasma, cerebrospinal fluid, interstitial space, glandular or renal tubular lumina, gall bladder, etc and each cell are all compartments which a drug may or may not enter or leave with different rate constants. In addition, binding to protein or other sequestration also is governed by characteristic rate processes. Consequently, it might be expected that the kinetics of absorption, distribution and elimination would be very complex and perhaps beyond analysis and mathematical description. Fortunately, the rates of distribution among the various tissues and myriad cells generally are not dispersed greatly, and most such processes are first-order. Thus, the kinetics behave as though the drug were being distributed among one, two or, at the most, a few compartments, and they are amenable to mathematical modeling. Like the volume of distribution (page 727), a pharmacokinetic compartment is fictive or virtual and may be difficult to define in precise anatomical terms. Therefore, a compartment is defined mainly by its pharmacokinetic parameters.

Open One-Compartment Model

In this model, the body is assumed to behave as though it were a single compartment, that is, as though there were no barriers to movement of a drug within the total body space and as though the final equilibrium distribution is attained instantaneously. In practice, the model adequately describes the pharmacokinetic behavior of a drug if the final equilibrium distribution is attained rapidly in comparison to the rates of absorption and elimination. The term *open* indicates that input and output (from any and all routes of administration and elimination, respectively) are unidirectional and that the one compartment (ie, body) is not within a confined space and hence does not come into chemical equilibrium with its external environment. In simple diagram, such an open one-compartment model is depicted in Fig 36-2. In the diagram, the compartment represents the entire body (excluding the lumina of the gastrointestinal tract, urinary tract, pulmonary alveoli, etc, which communicate with the open environment). The term, V_d is the *volume of distribution* (see page 727). However, V_d is not necessarily that of the body or even total body water; as noted on page 728, the volume of distribution, V_d, is a fictive one considered to be equal to fD/C_p (where f is the fraction absorbed, D is the dose and C_p is the plasma concentration) in which it hypothetically is assumed that the concentration is the same throughout the volume and is equal to the plasma concentration. In reality, concentration is not homogeneous throughout, but this cannot be determined from C_p alone (which simply averages all inputs and outputs); as long as distribution equilibrium is achieved rapidly, the kinetics as perceived through blood or urine concentrations are the same whether distribution is homogeneous or heterogeneous.

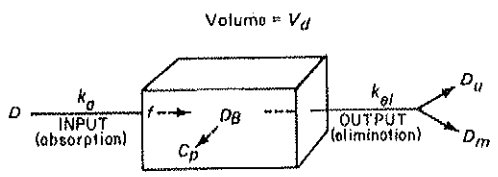


Fig 36-2. The open one-compartment pharmacokinetic model. An amount of drug, D_B, is absorbed from the administered dose, D, with a rate constant of k_a into a compartment with volume V_d and is distributed instantaneously to reach a plasma concentration C_p. V_d is obtained by dividing D_B by C_p. D_B = dose D times f, the fraction absorbed. Drug is eliminated from the compartment with a rate constant k_{el}. D_u is the amount excreted into urine, feces, expired air, sweat, milk, etc; D_m is the amount of drug metabolized.

In order to derive formulae to describe time-related changes in C_p, it is convenient to consider absorption and elimination separately, as though each were occurring in the absence of the other, then to add them algebraically to determine the total integral kinetics.

Absorption—If a drug is administered intravenously in a single, rapid injection, absorption is bypassed. The time for such injections is usually so short compared to other pharmacokinetic processes that it is customary to consider the peak plasma concentration and equilibrium distribution to occur instantaneously in one-compartment systems. This is depicted in panel A of Fig 36-3. In the model for the figure, there is no elimination and C_p remains constant once injection is accomplished. With constant intravenous infusion (panel B), C_p rises rectilinearly so long as infusion continues at a constant rate. With other routes of administration, absorption usually manifests first-order kinetics, since most drugs are absorbed by simple diffusion. Thus, the drug disappears exponentially from the site of administration (as from compartment 1 in Fig 36-1). The equation for the concentration of a drug in the plasma after a single extravascular dose of a drug, assuming no elimination takes place, is

$$C_p = C_p^{\infty} - C_p^{\infty} e^{-k_a t} \qquad \text{[units: wt} \cdot \text{vol}^{-1}\text{, etc]} \qquad (5)$$

where C_p is the concentration at time t, C_p^{∞} is the final concentration at "infinite" time and k_a is the absorption rate constant (units: time^{-1}). Absorption is characterized by a half-time equal to $0.693/k_a$. Bimolecular absorption processes, such as facilitated diffusion or active transport, also often show first-order kinetics, especially at drug concentrations well below those at which the carrier system will become saturated. At saturation, the kinetics become zero-order. Even the rate of dissolution of a drug approximates a first-order process, provided that the drug is soluble readily and diffuses rapidly. If the solubility and diffusibility are low, it will approximate a zero-order process so long as there is saturation around the solid phase. Some sustained-release dosage forms are designed to release drugs at a constant rate (zero-order) over long periods of time.

Absorption by the oral route rarely conforms to simple first-order kinetics. A drug is absorbed at different rates from the stomach and the three segments of the intestine, partly simultaneously and partly sequentially. Absorption from the stomach usually is quite slow compared to that from the small intestine, and it is sometimes so slow that a significant amount of drug appears in the blood only after the stomach contents are emptied. Thus, there may be a *lag* between the time of drug administration and the appearance of drug in the blood. That is, the curve describing the time-dependent rise in C_p does not pass through the origin. An example of lag in the absorption of pentobarbital is shown in Fig 36-4. Enteric-coated or other delayed-release dosage forms also cause lag. The mathematical formulation of lag

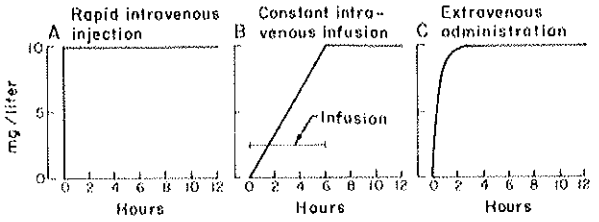


Fig 36-3. Time-concentration curves for injection (A), infusion (B) and extravenous (C) administration of drug in the one-compartment model. The volume of the compartment is 100 L (V_d = 100 L); the amount of drug administered in each instance is 1000 mg. Drug elimination has been set to zero, so that the time-concentration curve for each model of administration can be examined without the complication of simultaneous elimination (courtesy, Bigger,[1] adapted).

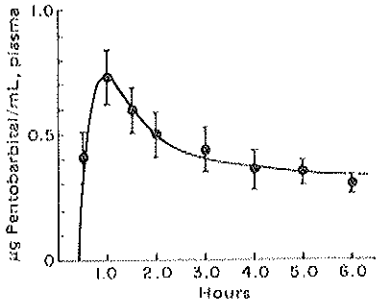


Fig 36-4. The time course of pentobarbital in the blood of a fasting human subject following the oral administration of 50 mg. The figure shows a lag-time of about 20 min, approximately the emptying time of the fasting stomach (courtesy, Dittert[2]).

will be deferred to the next section in connection with Eq 28. Factors affecting absorption are enumerated on page 713. Some changes in gastrointestinal conditions during the course of absorption are part of diurnal rhythms or are caused by the drug itself, which make it impossible to establish a steady basal state for description; others may result from emotionality, ingestion of foodstuffs, water, other drugs, etc, and can be controlled adequately for scientific purposes but may vary greatly in practical circumstances. Absorption by other routes is also subject to variability. Some drugs that are completely absorbed in normal patients may not be absorbed in persons with abnormal gastrointestinal function, as the result of genetic, pathological or surgical factors. Many drugs are not absorbed completely even when gastrointestinal function is optimal. Absorption can be limited by the physical state of the drug and by other substances in the dosage form. The amount of drug absorbed into the body (D_B) is related to the dose as follows:

$$D_B = fD \qquad \text{[units: wt]} \qquad (6)$$

where D_B is the amount absorbed (drug in the body), f is the fraction absorbed and D is the dose administered. The property of a drug to be absorbed from its dosage form is known as *bioavailability*, and f is the bioavailability factor. The bioavailability factor often is determined by comparison of the area under the concentration curve (AUC) of a given dose of drug given orally with that of the same dose given intravenously (see page 736).

Distribution—In the open one-compartment model, the body is treated as though it were a single compartment in which the absorbed drug is mixed instantaneously and homogeneously. Clearly, the assumption of instantaneous equilibrium establishes only an ideal mathematical boundary condition to facilitate pharmacokinetic calculations. At best, no drug could be equilibrated in less than one circulation time, and no drug has been shown to distribute so rapidly. However, for practical purposes, a distribution time of a few minutes is negligible compared to absorption and elimination times. Only water-soluble drugs of small molecular size which are confined completely to the extracellular space equilibrate rapidly enough to meet the requirements of the ideal one-compartment model, but, for clinical purposes, the one-compartment model is adequate to describe the pharmacokinetics of a large number of drugs.

Volume of Distribution and Distribution Coefficient—The hypothetical volume within which a drug is distributed is known as the *volume of distribution*, V_d. It may be calculated by dividing the amount of drug in the body, D_B, by the plasma concentration, C_p, where C_p is the concentration in plasma. It is important to note that C_p is usually the total concentration of unbound plus bound drug. Under real conditions, D_B and C_p vary with time, and computation

must be made in such a way as to eliminate the time variable. One such way is to extrapolate C_p to zero time (eg, see Figs 36-6 and 36-8), in which case

$$V_d = f\,D/C_p^0 \qquad (7)$$

where D is the dose administered, f is the bioavailability factor (fraction that reaches the systemic circulation) and C_p^0 is the plasma concentration at zero time, determined by extrapolation. When the drug is given intravenously, $D_B = D$.

Of course, V_d will vary with body weight, so that it needs to be normalized in a way that allows comparisons among individuals of different body weights. Such a normalized V_d is the *distribution coefficient*,* Δ', calculated by the equation

$$\Delta' = V_d/BW \qquad (8)$$

where BW is body weight. Units are usually mL/g or L/kg, and care must be taken to employ the appropriate units of weight, concentration and volume in Eqs 7 and 8. The notation Δ' is a more serviceable parameter than V_d and is the form of V_d usually found in tables of pharmacokinetic data, usually under the heading, "Volume of Distribution," rather then Δ'.

Although V_d and Δ' are derived as though the concentration was equal to C_p throughout the volume, concentration is, in fact, almost never homogeneous, and consequently V_d and Δ' are only imaginary (fictive, virtual) volumes. Factors that make for nonhomogeneous distribution are: binding to proteins, dissolution into body lipids, pH partition, active transport, electrochemical and Donnan distributions, etc. Even if C_p (free) rather than C_p (total) is used to calculate V_d, V_d would not represent a real space, because of these manifold factors that cause uneven distribution. Consequently, the principal utility of V_d or Δ' is not so much in permitting an estimation of where the drug is distributed but rather as a measure of the reservoir from which a drug is being delivered and/or cleared (see page 729 and Table II, page 731). However, with appropriate considerations, V_d or Δ' also may indicate the general ability of a drug to penetrate membranes, dissolve in fat or bind extensively to extravascular macromolecules.

Highly polar, poorly penetrant drugs tend to be confined mostly to the extracellular space; if these drugs are little bound to plasma proteins, they will have Δ's of about 0.3 mL/g, less if there is significant binding to plasma proteins. The lower limit to Δ' is about 0.04 mL/g, which approximately is equal to the plasma volume. Drugs that are distributed throughout body water and are not bound or concentrated have Δ's of approximately 0.7 mL/g, the Δ' of body water. Lipid-soluble drugs that are bound negligibly to plasma protein have Δ's that range usually from about 0.7 to 3–4 mL/g, depending upon water-lipid distribution coefficients. Some drugs that bind strongly to chromatin have Δ's that approach 1000 mL/g. However, many drugs combine penetrance, lipid solubility and protein binding in such proportions to make it difficult to interpret the meaning of Δ' without ancillary information.

Since, by definition, V_d varies reciprocally with C_p, it is essential to recognize that binding to plasma proteins, by increasing C_p, will decrease V_d. Despite this, plasma protein binding has no *real* effect on extravascular distribution. Since it is only the free form that moves among the spaces and tissues, it follows that alterations in plasma protein binding alone will not alter the extravascular (indeed, extraplasma) distribution. Only the calculated, fictive quantity, Δ', is affected. For example, nafcillin has a Δ' of 0.29 mL/g and is 90% bound to plasma proteins. If there was no protein binding, Δ' would equal 2.9 mL/g, a volume sufficiently larger than that of water, to suggest considerable extravascular binding. However, it is not the masking of the degree of extravascular distribution that is the source of difficulty when there is significant binding to plasma proteins, but rather because the extent of protein binding is not always constant. Both the quantity and binding properties (affinity and capacity) of human plasma proteins can vary in health, disease and the presence of other drugs (see pages 195 and 716). If the degree of binding of nafcillin to plasma proteins was to change to 50% as the result of hypoalbuminuria, Δ' would become nearly 0.48 mL/g. The Δ' of ampicillin, which is bound only to the extent of 18%, would not be affected so greatly. A further complication of binding to plasma proteins is occasioned when the degree of binding, and hence the magnitude of Δ', is dose-dependent. There are a number of known examples in which Δ' varies with the dose.

Elimination—Once a drug is absorbed, it is transported by the blood to the tissues, among which it is distributed, metabolized and/or excreted; all of these processes lower the plasma concentration of the drug. Each separate process ordinarily has first-order kinetics, and the overall change in plasma concentration is described by the linear combination (or algebraic addition) of the separate equations. In the one-compartment model, the kinetics of distribution are ignored, since distribution occurs so rapidly that distribution occurs before any practical blood-sampling or repetitive dosing occurs. Thus, after intravascular administration the plasma concentration, C_p, will fall exponentially according to Eq 1. Such an exponential elimination of theophylline, given intravenously, is shown in Fig 36-5. According to Eq 2, if the data of Fig 36-5 are plotted semilogarithmically, as in Fig 36-6, a straight line should result. Several derived data can be obtained from such a plot. Extrapolation to zero time (ie, the y intercept) gives C_p^0, the theoretical plasma concentration at time zero. It is a theoretical concentration, because neither injection nor distribution actually is instantaneous. Nevertheless, C_p^0 is a very practical figure. For example, from it may be derived the volume of distribution, V_d, simply by dividing the dose, D, by C_p^0 (see page 727). In the figure, $C_p^0 = 0.0115$ mg/mL, so that V_d is 43.5 L, or about 89% of the volume of total body water in a 70-kg adult. The plasma half-time, $t_{1/2}$, can be determined directly from the graph or from the elimination rate constant, k_{el}, by means of Eq 3. (Conversely, k_{el} could be derived from $t_{1/2}$, determined visually from a graph.) When determining k_{el} from the slope, it must be kept in mind that the log of the concentration must he used rather than the antilog that is

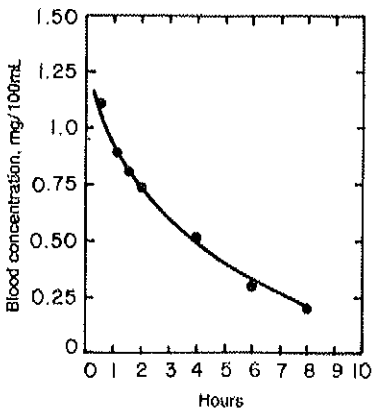


Fig 36-6. Elimination curve of average blood levels of theophylline in 11 human subjects after intravenous administration of 0.5 g aminophylline per 70 kg to each (courtesy, data, Truitt, *et al* [3]).

* Δ' is not to be confused with water-lipid distribution coefficients.

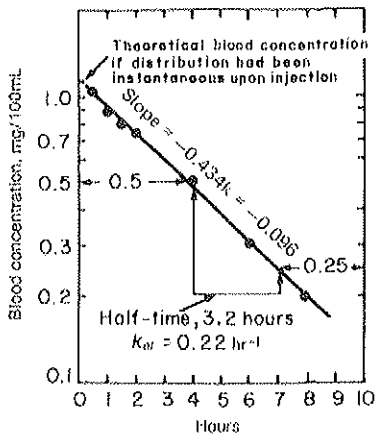


Fig 36-6. Semilog plot of the elimination curve in Fig 36-5. Note the log scale of the ordinate.

plotted on the log-scaled ordinate in the figure. In natural logarithms, the slope $(\ln C_{p1} - \ln C_{p2})/(t_2 - t_1)$ is equal to k_{el}; in decilogarithms, the slope $(\log C_{p1} - \log C_{p2})/(t_2 - t_1)$ is equal to $0.434k_{el}$. From the figure, k_{el} is found to be 0.22 hr^{-1}. However, this is an *instantaneous* rate and is not the same as the fraction that disappears over a finite interval, eg, 1 hr. Nevertheless, it is sometimes convenient to use k_{el} for calculation of the amount of drug lost per unit time, eg, for the calculation of maintenance doses during chronic drug therapy. If the time interval under consideration is very short (at least 1/3) as compared to the half-life of the drug, k_{el} may be considered to be an indication of the percentage of the drug lost during the time interval. For example, if the half-life of a drug is 6.93 hr, k_{el} would equal 0.1 hr^{-1}, and approximately 10% of the drug would be lost in 1 hr. Thus, if 100 mg of drug were present at the beginning of the hour, 10 mg would be lost by the end of the hour. (The exact amount of drug lost is determined by use of Eq 2 and is found to be 9.51 mg in 1 hr, a difference of about 5% from the approximated value.) If such an approximation was used in the example given in Fig 36-6 (1 hr is a little less than 1/3 of the half-life) and the amount of drug present is 100 mg, the figures comparable to the illustration above would be 22 versus 19.7 mg, respectively, almost a 12% difference.

Within the group of 11 subjects in the above study, there were considerable differences in k_{el} or $t_{1/2}$ among the members. One cannot overemphasize the caveat not to take too literally the average half-life data found in various tables or other literature but rather to assume a probability that the half-life in a given patient may depart considerably from that average value. The half-lives of some drugs vary over a wide range even in normal individuals. The half-life of amitriptyline, a drug with a complex metabolic and excretory elimination, varies nearly tenfold; even the half-life of penicillin, a drug with a simple excretory elimination, varies twofold. In persons with hepatic or renal failure, the published mean half-life data may not even be in a range applicable to such persons.

The half-life also may vary widely from species to species; for example, in man, the half-life of sulfaethidole is about 8 hr, whereas in cattle it is less than 2 hr. Half-lives also vary considerably even among congeneric drugs, as may be seen with the sulfonamides shown in Table I.

The biological half-life must not be confused with the time for the response to decline by 50%, since dose, the requirement for a threshold concentration, latency of response and other factors may cause a nonparallelism between blood concentration and intensity of response. In fact, because the relationship between effect and plasma concentration is usually logarithmic, effect tends to decline in a linear, not loglinear, fashion.

Table I—The Approximate Biologic Half-Life in Man of Several Sulfonamides[4]

Drug	$t_{1/2}$ (hr)
Sulfamethylthiadiazole	2
Sulfaethidole	8
Sulfisoxazole	8
Sulfamethoxypyridazine	34

In Fig 36-2, the rate constant for elimination is designated k_{el}, without reference to the mode or route of elimination. However, k_{el} may be a compound constant equal to the sum of the rate constants that define the various simultaneous (ie, parallel) contributory processes, such as biotransformation, renal excretion, biliary secretion, etc. Thus, the compound or overall constant is $K = k_1 + k_2 + k_3 \ldots k_n$ where k_1, $k_2 \ldots k_n$ are the rate constants of the separate contributory processes. Consider the case in which a drug simultaneously is biotransformed and excreted unchanged in the urine. The initial concentration, C_p^0, therefore will be diminished by both $C_p^0 e^{-k_m t}$ and $C_p^0 e^{-k_u t}$, where m designates metabolism and u renal excretion. (In some notations, k_u is designated k_r, k_{10}, k_3 or k_e.) Therefore, Eq 2 adapted for the two processes becomes

$$\log C_p = \log C_p^0 - 0.434k_m t - 0.434k_u t$$
$$= \log C_p^0 - 0.434(k_m + k_u)t \qquad \text{[no units]} \quad (9)$$

Thus, $k_m + k_u$ combine to make a single constant, which is the overall elimination rate constant. In order to identify it as a compound, or overall, constant, it is sometimes designated as K, rather than k_{el}.

Clearance and Routes of Elimination—The half-life of a drug is a useful pharmacokinetic parameter. Since half-life is expressed in units of time, it is an easily understood, concise indication of the rate of disappearance or accumulation of a drug. Further, it is used to estimate the time necessary to attain a new steady state whenever a steady state is altered by a change in the factors determining dose regimen, namely, drug dose, bioavailability, the dose interval, rate of elimination and volume of distribution (see page 740). However, the elimination half-life of a drug is a complex function of drug distribution, biotransformation and elimination. A more direct expression of the rate of drug elimination is drug clearance.

Clearance is the rate of removal of a drug or other substance from the body, expressed as the *in vivo* volume equivalent of the substance being removed per unit time. In order to illustrate the concept, assume that drug D is being eliminated from the body at a rate of 0.1% per minute. The absolute amount of drug that was eliminated would therefore be equal to 0.1% D_B per minute. Since D_B is distributed as though it were in a volume V_d (volume of distribution, page 740), one can calculate the fictive volume equivalent of the amount of drug lost per minute, which in this instance would be 0.1% V_d/min. Since the relative rate of loss, 0.1%/min or 0.001/min is, in fact, k_{el}, it may be seen that

$$Cl_t = k_{el} \times V_d \qquad \text{[vol/unit time]} \quad (10)$$

where Cl_t is total body clearance. It may be expressed in units of mL/min, L/hr or mL/kg/min; the reader must be alert to the units in which given clearance data are expressed. It must be emphasized that clearance is a hypothetical or fictive quantity, since the body rarely clears a drug completely from a specific volume of body fluid. Only when elimination is flow-limited is the blood that passes through the eliminating organs(s) totally cleared, so that the effluent blood is essentially devoid of drug; in such an in-

stance the clearance approximates the rate of blood flow. If the concentration in the effluent blood were to be only 0.5 of the affluent blood, the clearance would be said to be 0.5 that of the blood flow.

Although clearance is the dV/dt equivalent of dD_B/dt, or the volume equivalent of the drug lost per unit time, the hypothetical volume cannot be regarded as also having been eliminated. Just as the depleted effluent blood from the eliminating organ is returned to the systemic circulation to mix with all the blood and as drug is redistributed and re-equilibrated among the vascular and extravascular components of V_d, the fictive volume that is "cleared" remains a part of V_d, so that the only change that is effected is one in concentration, of which C_p is the index. Since V_d and k_{el} are both constant, it follows that Cl_t is also constant.

The concept of clearance can be applied to the whole body or to specific organs. The former application is a convenient way to indicate overall drug elimination; the latter application is used to indicate the contribution of a specific organ to drug disappearance.

Total Systemic (Whole Body) Clearance—Total body clearance is the sum of all the separate clearances that contribute to drug elimination, ie, $Cl_{tot} = Cl_{metab} + Cl_{renal}$, etc. It is essential that k_{el} be expressed in the same time units as are used in clearance (usually min). In Eq 10, dividing by 60 converts k_{el} in hr^{-1} to min^{-1}, so that clearance can be expressed in mL/min. Whole-body clearance in a one-compartment system is also equal to dose divided by the area under the curve:

$$Cl_{tot} = D/AUC_0 \qquad [mL \cdot min^{-1}] \qquad (11)$$

where AUC_0 is the area under the curve (AUC), discussed on page 736. The determination of Cl_{tot} in the two-compartment system is discussed on page 739.

Renal Excretion and Clearance—The principles of renal excretion and clearance have been used for approximately 50 years as tools for studying renal physiology and pathology and hence were adapted early to pharmacokinetics. Consequently, renal clearance of drugs is a classic illustration of the general subject of clearance. As discussed in Chapter 35, all drugs are filtered in the glomerulus and some also are secreted into the urine by renal tubular cells; there is also resorption of drugs from the tubular luminal fluid back into the blood as the fluid passes along the tubule. Glomerular filtration is the passage through the glomerular vascular endothelium of the plasma fluid and all solutes therein small enough to pass through the pores; that is, it is the filtration of water and all micromolecular solutes. Thus, it is independent of the presence of drug and is a function of the filtration pressure (which relates to blood pressure) and the mean transit time across the glomerular capillaries. The rate of filtration is known as the *glomerular filtration rate*, *GFR*, and it has the units of vol/min (usually mL/min). In turn, the transit time is determined by the rate of flow of blood through the glomeruli; this rate of blood flow is known as the *renal plasma flow*, *RPF*. Since only a fraction of the plasma is filtered during passage through the glomerulus, it is useful to designate this fraction as the filtration fraction, *FF*, where $FF = GFR/RPF$. The average renal *plasma* (not blood) flow in the adult human male is approximately 600–700 mL/min, and the *GFR* is approximately 100–125 mL/min (of which 99% of the water is resorbed and returned to the blood); thus, the filtration fraction is approximately 0.2.

Under basal conditions, the *GFR* is roughly constant in time. Therefore, the only major variable that determines the rate of filtration of free drug is the concentration of drug in the plasma. Thus,

$$F = C_{pf} \cdot GFR \qquad [\text{units: } mL \cdot min^{-1}] \qquad (12)$$

where F is the filtration rate of the drug, usually in units of mg/min, and C_{pf} is the amount of *free* drug in the plasma. If the drug is unbound, $C_{pf} = C_p$. If the drug is bound to plasma protein, then

$$F = [C_p(1 - p)] \cdot GFR \qquad [mL \cdot min^{-1}] \qquad (13)$$

where p is the fraction bound to plasma protein.

The *GFR* may be determined by the steady-state rate of excretion of any nonbound chemical substance that is not secreted subsequently and/or resorbed by the renal tubules, so that the amount of substance which appears in the urine is all of that which was filtered and no more. Two such substances are *creatinine* and *inulin*. With creatinine, the endogenous plasma levels are nearly constant, and thus creatinine lends itself readily to the determination of *GFR*. Either inulin or creatinine may be given by constant intravenous infusion; usually, creatinine is used. However, it is not customary to express the *GFR* of creatinine or of drugs as F, in terms of mg/min, but rather in terms of *clearance*. As discussed above, clearance is a hypothetical volume of plasma which, if completely cleared of its content of drug in unit time, would be equivalent to the amount of drug that disappears in unit time. In the instance of filtration, it is easy to visualize clearance as that volume filtered/min, since the filtered volume actually is separated physically from the blood. Thus, the *creatinine clearance*, or *GFR*, is equal to the total amount of creatinine found in the urine (equal to urine concentration times urine volume) divided by the plasma concentration.

The general concept of clearance can be applied to the kidney according to the equation

$$Cl_{ren} = \frac{\bar{C}_u V}{\bar{C}_p t} \qquad [mL \cdot min^{-1}] \qquad (14)$$

where Cl_{ren} is renal clearance, $\bar{C}_u$ is concentration in mg/mL of drug in urine collected during time t, V is urine volume in mL generated in time t (min) and $\bar{C}_p$ is the mean concentration (during the collection interval, t) of drug in the plasma in mg/mL; the units are thus mL/min. Urine is collected from the bladder by catheter or by voiding. At the beginning of the collection interval (time from last voiding) both $\bar{C}_p$ and $\bar{C}_u$ are higher than at the end. Consequently, $\bar{C}_p$ must be calculated; $\bar{C}_u$ is automatically the mean of the instantaneous collecting duct concentrations. Equation 14 is valid whether the drug is "cleared" by filtration or by tubular secretion and whether or not tubular resorption occurs. If the drug is protein-bound, the formula becomes

$$Cl_{ren(corr)} = \frac{\bar{C}_u V}{\bar{C}_p t(1 - p)} \qquad [mL/min] \qquad (15)$$

where $Cl_{ren(corr)}$ is the corrected renal clearance.

The ratio between Cl_{ren} and Cl_{creat}, Cl_{ren}/Cl_{creat} (or Cl_{ren}/Cl_{inulin}), is known as the *clearance ratio*. If the drug is protein-bound and the *corrected clearance* is used, the ratio $Cl_{ren(corr)}/Cl_{creat}$ is known as the *excretion ratio*.

If an unbound drug is filtered only and not resorbed, the excretion ratio will be 1 and the clearance about 125 mL/min; if the drug subsequently is resorbed, the excretion ratio will be less than 1 and the clearance will lie between 125 and 1 mL/min, the values depending upon the degree of resorption. A clearance of 1 mL/min suggests distribution and elimination like those of water. If there is tubular secretion (plus obligatory filtration), the excretion ratio may exceed 1, and the clearance could be as high as 600–700 mL/min, depending upon the extent of tubular secretion and resorption. *Para*-aminohippuric acid (PAHA) is not bound to plasma protein, is not tubularly resorbed and is secreted so fast by the renal tubules that the plasma passing through the kidney is 90% cleared of PAHA. Thus Cl_{PAHA} is

equal to 0.90 *RPF*. This is called the *effective renal plasma flow, ERPF*. The excretion ratio of PAHA is about 5 to 6.

Eq 14 can be rearranged so that

$$\frac{C_u}{t} = \frac{C_p Cl_{ren}}{v} \qquad [\text{wt} \cdot \text{vol}^{-1} \cdot \text{min}^{-1}] \qquad (16)$$

Thus, it may be seen that the concentration of drug in newly formed urine is directly proportional to the plasma concentration. Since the plasma concentration falls exponentially during the collection interval, t, it follows that the instantaneous urine concentration in the collecting ducts, likewise, must fall exponentially and hence the rate of fall can be expressed by a first-order rate constant, k_u. This constant relates to renal clearance as follows:

$$k_u = \frac{Cl_{ren}}{V_d} \qquad [\text{min}^{-1}] \qquad (17)$$

The excretory rate constant may be simple, as with a drug like creatinine, or compound, as with a drug that is secreted tubularly and/or resorbed.

The overall renal elimination constant, k_r, is defined by

$$k_r = k_g + k_{ts} - k_{tr} \qquad [\text{min}^{-1}] \qquad (18)$$

where k_g is the constant for glomerular filtration, k_{ts} for tubular secretion and k_{tr} for tubular resorption. Although k_r might be thought to be the same as k_u on page 729, in practice it is not, because clearance data are obtained from time-averaged concentrations and cannot provide instantaneous rates. However, creatinine-derived k_r is close to the instantaneous k_u at the midpoint of the collection period.

By combining Eqs 3 and 17 and assuming that there is no other route of elimination,

$$t_{1/2} = 0.693 \frac{V_d}{Cl_{ren}} \qquad [\text{time}] \qquad (19)$$

The units of time must be the same for both $t_{1/2}$ and Cl_{ren}. The equation enables the calculation of some thought-provoking information about the biological half-lives of nonmetabolized drugs of different excretion profiles and volumes of distribution. Approximate hypothetical half-lives of drugs of different volumes of distribution and renal clearance are shown in Table II. The drugs are assumed to be eliminated only by renal excretion. A volume of distribution of 50 L is that of total body water, 15 L is that of extracellular water and 50,000 L is that of a drug strongly bound in the tissues. Because of biotransformations, few drugs have half-lives longer than 1 yr. However, a few radioopaque iodine-containing diagnostic agents are so tightly bound that their half-lives exceed 1 yr. At the other extreme, a half-life of 15 min by renal elimination is uncommon, because few drugs that are totally cleared have volumes of distribution as small as that of extracellular water. However, the half-life of penicillin G is about 30 min.

Although data from collected urine cannot provide instantaneous rates, it does allow the calculation of the plasma half-life. The instantaneous excretion rate, dD_u/dt (where D_u is the amount of drug in urine), is directly proportional to the body burden, D_B, such that

$$dD_u/dt = k_u D_B \qquad [\text{wt} \cdot \text{min}^{-1}] \qquad (20)$$

But, D_B is falling exponentially with a rate constant k, so that $D_B = D_B^0 e^{-kt}$; therefore, $dA/dt = k_u D_B^0 e^{-kt}$. It follows that the slope of a plot of the log of the excretion rate versus time will have a slope of $-0.434k$, analogous to Eq 2 (adapted to total content rather than concentration). The y intercept of such a plot is $\log k_{el} = D_B^0$, where D_B^0 is the amount of drug in the body at zero time. However, data on excretion rates require renal catheterization and are subject to considerable error. An alternative, usually more accurate, method of estimating k from urine concentration is to employ the cumulative amount excreted. In this method,

$$D_u = \frac{D_B k_u}{k}(1 - e^{-kt}) \qquad [\text{wt}] \qquad (21)$$

Since k_u/k expresses the proportion of D_B being transferred to the urine, $D_B^0 k_u/k$ represents the total amount of drug excreted, or D_u^∞, where ∞ designates infinite time. Eq 21 in log form, with the above substitution and transposition, becomes

$$\log (D_u^\infty - D_u) = \log D_B^0 \frac{k_u}{k} - 0.434kt$$
$$= \log D_u^\infty - 0.434kt \qquad [\text{no units}] \qquad (22)$$

The slope of the plot against time is also $-0.434k$ and $(D^\infty - D_u)$ is the amount of drug that remains in the body. The equation applies if the drug is administered intravascularly. This is known as the sigma minus method (sigma for the integral D_u^∞ and minus for the $-D_u$). Urine needs to be collected for only 3 or 4 half-lives in order for the semilog plot to yield a reliable slope and $t_{1/2}$. The method is useful especially when plasma concentrations are low.

Hepatic Clearance—The concept of hepatic clearance is like that of renal clearance, and hepatic clearance is likewise a hypothetical volume of blood per min imagined to be totally cleared of drug during passage through the liver. Unlike renal clearance, the input is both portal venous and hepatic arterial blood and the output is both hepatic venous blood and bile, rather than arterial blood and urine, respectively. Portal venous blood and bile cannot be sampled readily, so that the concepts involved in hepatic clearance serve better to provide a model for understanding the role of the liver in pharmacokinetics than a clinical methodology for its direct measurement.

Although the mathematical treatment of hepatic clearance has been developed for steady-state conditions, rather than for exponentially falling drug concentrations in the inputs and outputs to the liver, the subject is appropriate at this place, in conjunction with other clearances.

The *hepatic clearance*, Cl_H, can be defined by the equation

$$Cl_H = HBF\left(\frac{C_{ap} - C_v}{C_{ap}}\right) = HBF \cdot E \qquad [\text{mL} \cdot \text{min}^{-1}] \qquad (23)$$

Table II—Hypothetical Half-Lives of Drugs of Differing Volumes of Distribution and Clearances

Drug No.	Distribution	V_d, L	Renal Disposition	Clearance, mL/min	Half-Life
1	Total body water	50	Filtered and resorbed with water	1	24 days
2	Total body water	50	Filtered, no resorption	125	4.67 hr
3	Total body water	50	Tubular secretion, total clearance	700	50 min
4	Extracellular water	15	Tubular secretion, total clearance	700	15 min
5	Strongly bound in tissues	50,000	Filtered and resorbed with water	1	66 yr
6	Strongly bound in tissues	50,000	Tubular secretion, total clearance	700	35 days

where HBF is the total hepatic blood flow, C_{ap} the hypothetical mean of mixed hepatic arterial and portal venous concentrations and C_v is the hepatic venous concentration. The ratio, $(C_{ap} - C_v)/C_{ap}$, is the *extraction ratio*, E. Unlike glomerular filtration, there is an upper limit to the absolute quantity of drug that can be cleared and hence to the extraction ratio. Extraction is flow-limited only so long as the biotransforming enzyme system is not approaching saturation. The maximal clearance in the presence of normal blood flow has been called the *total intrinsic clearance*, Cl_{intr}. The extraction ratio expressed in terms of Cl_{intr} is

$$E = \frac{Cl_{intr}}{HBF + Cl_{intr}} \qquad [\text{no units}] \tag{24}$$

which may be substituted into Eq 23, to yield

$$Cl_H = HBF\left(\frac{Cl_{intr}}{HBF + Cl_{intr}}\right) = HBF \cdot E \qquad [\text{mL} \cdot \text{min}^{-1}] \tag{25}$$

The intrinsic clearance becomes

$$Cl_{intr} = \frac{HBF \cdot E}{1 - E} \qquad [\text{mL} \cdot \text{min}^{-1}] \tag{26}$$

Cl_{intr} is thus somewhat analogous to V_{max}/K_m in enzyme kinetics.

Eqs 23 through 26 emphasize that hepatic clearance and extraction are functions both of hepatic blood flow and the capacity of hepatic enzymes to biotransform (or secrete into bile) the drug that is delivered. In order to appreciate the relative dependencies on Cl_{intr} and HBF, various assumed values may be substituted into the equations. What will be found is that the larger the Cl_{intr}, the more Cl_H tends to be flow-limited (ie, dependent upon the rate of delivery of blood), whereas when Cl_{intr} is small, Cl_H is metabolism-limited. At constant blood flow with a drug in which elimination is predominately hepatic, when intrinsic clearance and hence extraction ratios are small, a significant change in intrinsic clearance will be accompanied by a significant change in $t_{1/2}$; when intrinsic clearance is high, a significant change may be accompanied by a small, often insignificant, change in $t_{1/2}$ but a significant decrease in bioavailability. In the latter instance, $t_{1/2}$ is determined mostly by the fraction of the cardiac output that passes through the liver. Figures illustrating these features, an excellent discussion of hepatic clearance and various models of hepatic elimination are available,[b] as well as a treatment of the effect of binding of drug to plasma protein. Binding to plasma protein limits clearance when intrinsic clearance is low but not when it is high.

Although the determination of Cl_{intr} is too involved for routine investigative purposes, it may be estimated according to the equation

$$Cl_{intr} = \frac{\left(1 - \frac{D_{un}}{fD}\right)D}{AUC_0} \qquad [\text{mL} \cdot \text{min}^{-1}] \tag{27}$$

where D_{un} is the total quantity of drug excreted unchanged, f is the fraction absorbed, D is the dose administered and AUC_0 is the total area under the blood concentration-time curve after intravenous administration. The meaning of AUC will be discussed later (page 737).

Some drugs may be used to illustrate some of the points emphasized by the model. For example, at blood concentrations of ethanol above 0.02–0.04%, the hepatic alcohol dehydrogenase system is saturated, and hence hepatic blood flow will have little effect on Cl_H of ethanol above the concentration indicated. This implies that liver disease or injury will not much affect the rate that ethanol is cleared from the blood, a fact of some forensic importance. The hepatic biotransformations of pentobarbital and phenytoin are relatively slow, ie, Cl_{intr} are low; consequently, the induction of hepatic cytochrome P450 will increase Cl_H, almost in proportion to the degree of induction, and the $t_{1/2}$ will be shortened accordingly. The hepatic biotransformation of lidocaine is extremely rapid, ie, the Cl_{intr} is very high, so that Cl_H is limited by hepatic blood flow. This means that by the oral route, in which all of the absorbed drug obligatorily passes through the liver, only very small amounts will survive the pass through the liver into the systemic circulation. This nearly total clearance as the drug passes through the liver into the rest of the body is known as the *first-pass effect*. The clinical significance of the first-pass effect is discussed in Chapter 37. The flow-limitation in the hepatic metabolism of lidocaine also means that in congestive heart failure or shock, in which hepatic blood flow is diminished, the rate of biotransformation will decrease and $t_{1/2}$ will increase.

Biliary secretion contributes to hepatic clearance and hence is included in the above pharmacokinetic considerations. However, drugs that are excreted intact or in a form from which the drug can be sequestered in the intestines and subsequently resorbed (enterohepatic recirculation) may have complex pharmacokinetics if the rate of biliary secretion is an appreciable fraction of the hepatic clearance and if the enterohepatic reservoir is large.

Other Routes and Clearances—The kidney and liver are usually the major organs in the elimination of drugs, and all other routes combined often contribute negligibly. However, with the volatile anesthetics, pulmonary clearance is the major route, and pulmonary clearance becomes dominant; pulmonary clearance of gases is flow-limited. With some drugs, mammary secretion is appreciable, and the presence of drug in milk may present hazards to nursing children; however, pharmacokinetics in the mother usually is not affected by lactation. Salivary secretion is too small to affect systemic pharmacokinetics, but the concentration of drug in saliva usually parallels that in plasma, so that, with some drugs, it is possible to follow systemic pharmacokinetics by sampling saliva.

Absorption Plus Elimination—The kinetics of absorption and disposition must now be put together to define the time-related curve which describes the plasma concentration of a drug administered *extra*vascularly. The curve is determined by the algebraic sum of all processes involved in absorption, distribution and elimination. Since disposition (distribution plus elimination) begins as soon as the drug enters the blood stream, the plasma concentration reflects all these processes from the outset.

The time course of the plasma concentration of a drug in a one-compartment body can be obtained by combining algebraically Eqs 1 and 5, with appropriate rate constants, and substituting fD/V_d for C_p^∞. When the equations for absorption and elimination are thus combined

$$C_p = \frac{fDk_a}{V_d(k_a + k_{el})}\left(e^{-k_{el}t} - e^{-k_a t}\right) \qquad [\text{wt} \cdot \text{vol}^{-1}] \tag{28}$$

where f is the fraction absorbed, D is the dose, etc.

This equation simplifies to

$$C_p = C_p^0 e^{-k_{el}t} - C_p^0 e^{-k_a t} \qquad [\text{wt} \cdot \text{vol}^{-1}] \tag{29}$$

where C_p^0 and the C_p^∞ of Eq 5 are the same, since they both represent all of dose, D, distributed in V_d. If there is a lag, the t-factor in the exponents of e should be $t - t_1$, where t_1 is the lag time. Figure 36-7 shows a plot of the plasma concentration for each of absorption and elimination separately and when the two are combined.

In Fig 36-7 the parameters of absorption and elimination were assumed, in order to construct the figure. In practice, drug concentration-time data are obtained empirically, and the parameters are obtained from a semilog plot of the data,

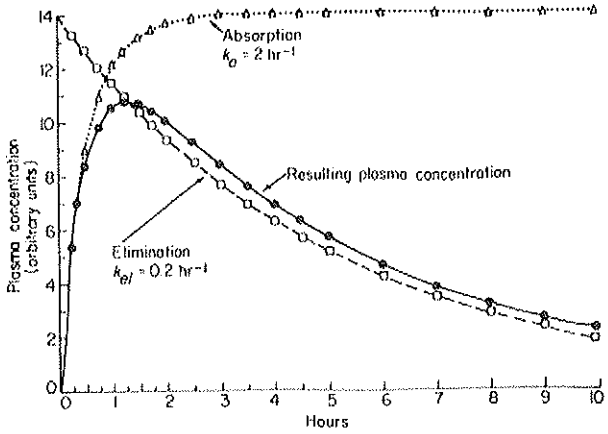


Fig 36-7. Time course of the plasma concentration of a hypothetical drug with simple first-order absorption and elimination kinetics. The rate constants are shown in the figure. The half-time for elimination is 3.47 hr.

as in Fig 36-8. The rising phase of the plot is not log-linear, since that which is added by absorption is diminished by elimination. Only after absorption is complete does the plot become log-linear, since now there is no opposing process at work against the mono-exponential decline in concentration. The time at which absorption essentially is complete is called the *absorption time* and is detected as that time at which the plot becomes log-linear. However, prior to the absorption time, the concentration at the site of deposition becomes equal to that in plasma. This is called the *equilibrium time*. It is also the *peak-time* for plasma concentration. Because of the interplay of physicochemical and active transport factors that affect the distribution of a drug, true chemical equilibrium is not reached necessarily at the pharmacokinetic equilibrium point. The log-linear line described by the elimination phase, when back-extrapolated to the y-axis, yields a theoretical C_p^0, just as with intravascular injection, and $V_{d(extrap)}$ can be calculated accordingly. Furthermore, the slope of the log-linear elimination segment of the semilog plot is equal to $-0.434k_{el}$, as with intravascular injection. The absorption rate constant, k_a, also can be obtained from the plot, if the empirical curve is subtracted from the back-extrapolated elimination line. This is done by subtracting the real values for C_p (C_1, C_2, etc) at various times during the absorption phase from the extrapolated values for C_p, designated C', on the back-extrapolated elimination line. It must be remembered that the antilog and not the log of C must be used if *log* C is plotted in Cartesian coordinates. This method of dissecting a compound function into its separate components is known as the *method of residuals*, or *back-feathering*. The back-feathered *absorption* line is the dashed line; its slope is negative, as though it were being seen from the site of administration.

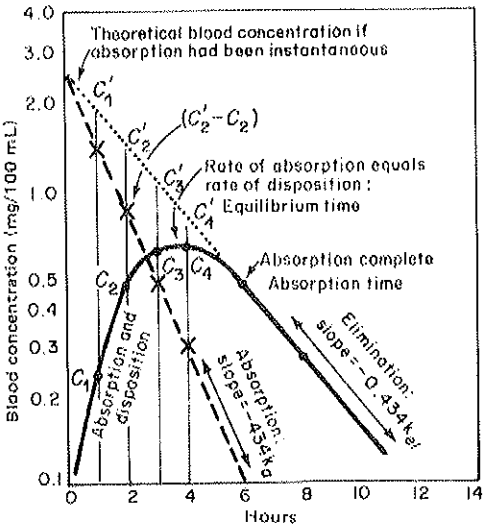


Fig 36-8. Kinetics of absorption and disposition of theophylline in the human subject after oral administration of 0.5 g of aminophylline per 70 kg. Blood concentration is plotted on a log scale (courtesy, data, Truitt, *et al* [3]).

The *peak concentration*, *time of peak concentration* and *duration of action* are affected by various factors, some of which are discussed below.

Peak Concentration—That the peak concentration should vary with the dose is self-evident; according to Eq 28, it is directly proportional to the dose (assuming that absorption and elimination are first-order processes). Figure 36-9 shows how peak concentration varies directly with dose. Note that the time of peak concentration is the same for all doses; this independence of peak time from dose is approximately true in all multicompartment systems. Departures from the generalization occur especially when the rate of absorption or elimination is different at high from those at low concentration; ie, when it is dose-dependent (see page 744).

Time of Peak Concentration—The time of peak concentration must not be confused with the time of peak effect. Effect often lags behind plasma concentration, sometimes because the tissue concentration at the point of action has not yet reached its peak and sometimes because a response may have a considerable latency. The latency of effect of reserpine or phenytoin (in its anticonvulsant effect) is measured in hours to days. Occasionally, the time of peak effect may precede the time of peak concentration because of a reflex or other compensatory process which limits effect before the concentration becomes maximal. This is often true with oral administration of ethanol or ephedrine. Both the peak concentration and time of peak concentration are considerably affected by the rate constants for absorption and elimination. In Fig 36-10, the effect of differences in absorption rate is shown indicating that the higher the absorption rate, the higher the peak concentration and the earlier the time of peak concentration. Figure 36-11 shows the effect of differences in rate of elimination depicting that the higher the elimination rate, the lower the peak concentration but the earlier the time of peak concentration. The two effects of absorption rate and elimination rate can be

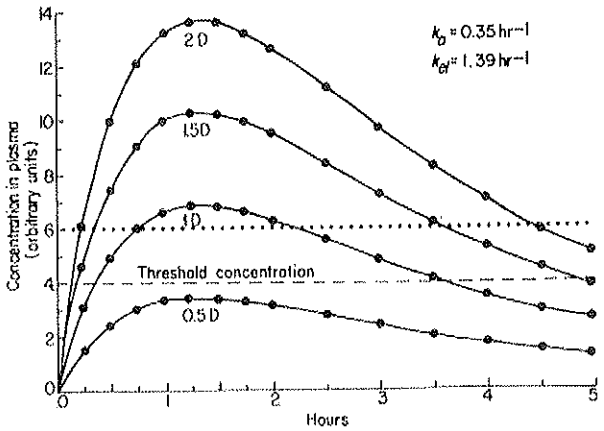


Fig 36-9. The effect of the size of the dose of a drug on the peak concentration, time of peak concentration and duration of action. The data are calculated from a one-compartment model.

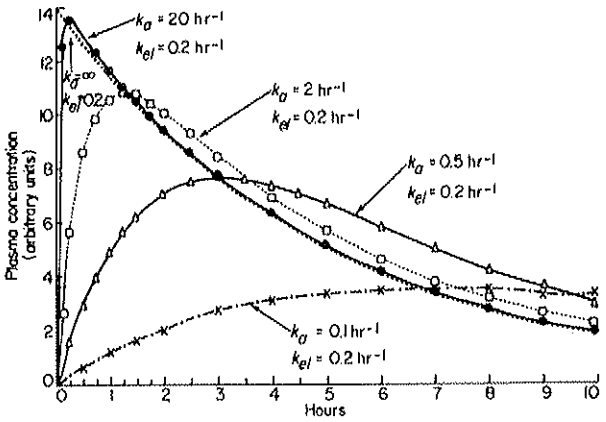


Fig 36-10. The effect of differences in the rate of absorption of drugs on the peak concentration, time of peak concentration and sojourn in the body. The rate of elimination is the same for all curves. The dotted line ($k_a = \infty$) is approximately what the concentration curve would be, had the drug been given intravenously. The data were calculated from a one-compartment model.

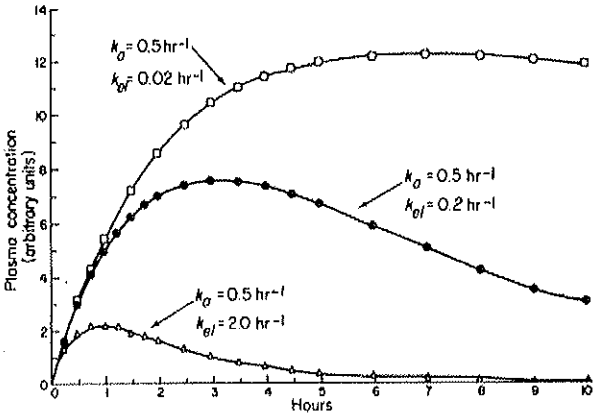


Fig 36-11. The effect of differences in the rate of elimination of drugs on the peak concentration, time of peak concentration and sojourn in the body. The rate of absorption is the same for all curves. The data were calculated from a one-compartment model.

treated as a single phenomenon if the ratio of k_a/k_{el} is considered rather than the separate rate constants (Fig 36-12).

The effects illustrated in Figs 36-10 to 36-12 have certain clinical implications:

1. *Differences in the rate of absorption are of more significance for slowly than for rapidly absorbed drugs.* In Fig 36-10, the peak blood levels are achieved when $k_a = 2\ hr^{-1}$ ($t_{1/2} = 0.35$ hr) is only 13% lower than when $k_a = 20\ hr^{-1}$ ($t_{1/2} = 0.035$ hr), but the difference in the level when $k_a = 0.1\ hr^{-1}$ ($t_{1/2} = 6.93$ hr) is 49% lower than that when $k_a = 0.5\ hr^{-1}$ ($t_{1/2} = 1.39$ hr), even though in the latter the rate difference was less than in the former comparison. It is thus apparent that differences in the release rates among different products of the same drug, or that differences in gastrointestinal motility, blood flow, etc, may be important, depending upon k_a/k_{el}. This point has a special relevance to sustained-release and depot formulations. With a number of drugs, especially among the anorectic drugs, the dose with a sustained-release form often is approximately the same as that of a rapid-release form; thus, the former has a long duration in the body but yields low blood levels when used in a single dose. Except with the initial dose, the differences are of lesser importance in a multiple-dose regimen. Small differences in the rate of absorption of rapidly absorbed drugs are usually of minor significance.

2. *When the rate of absorption is rapid relative to that of elimination, differences in the rate of elimination do not greatly affect the peak concentration consequent to a single dose* (compare top two curves of Fig 36-12). Thus, in such instances, the peak concentration is relatively insensitive to normal variations in the rate of elimination. Consequently, with such a drug, the size of the initial dose in a multiple-dose regimen often may not need to be diminished in the presence of renal or hepatic impairment; however, subsequent doses require adjustment.

3. *A change in the time of peak concentration or of peak effect is usually an indication of a change in one of or both k_a and k_{el}.*

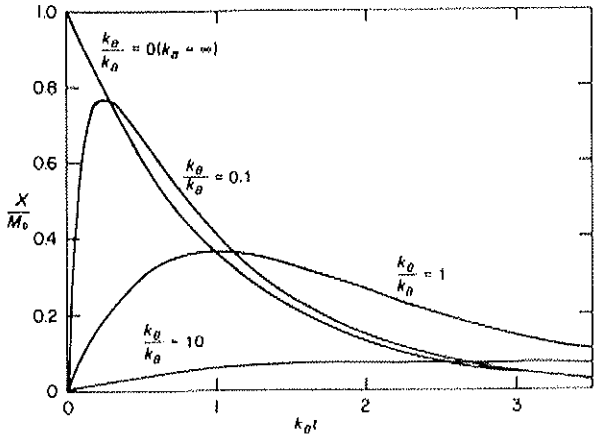


Fig 36-12. The effect of differences in the ratio, k_{el}/k_a (k_a/k_a in diagram), on the peak concentration, time of peak concentration and sojourn in the body. The ordinate, X/M_0, actually represent the fraction of a dose that is in the body, but they are directly proportional to concentration and thus serve to represent concentration. The abscissa can be converted to time by dividing by k_e, the elimination rate constant (courtesy, Goldstein, *et al* [6]).

Duration of Action—The duration of action of a drug is related to its pharmacokinetics in a rather complicated way. It is usually shorter than the sojourn of the drug in the body, because a threshold, or minimal effective, concentration must be reached before the effect occurs (see Fig 36-9), and the effect usually ceases when the plasma concentration falls below the threshold level. In a one-compartment system, duration of action tends to be proportional to log-dose. In a two-compartment system, it tends to be proportional to log-dose only when the site of action is in the central compartment and the effective concentrations (minimum to maximum) are entirely within the concentrations found during the elimination phase. In Fig 36-9, the duration of action is 3.25 hr with dose D, 4.6 hr with $1.5D$ and 5.4 hr with $2D$; were the threshold at 6 (dotted line), instead of 4 (dashed line), the respective durations would have been 1.5, 3.25 and 4.25 hr. Although the example in which the threshold is 6 provides that the duration of action would be disproportionately prolonged as the dose is increased, the contrary is seen when the threshold is 4. Consequently, *increasing the dosage is usually not a feasible way of increasing the duration of action* and toxic concentrations are often reached more predictably than duration is prolonged.

With a few drugs, there is no mathematically definable relationship between duration of action and persistence of the plasma concentration. With reserpine, for example, the effect outlasts the sojourn of the drug, because of the depletion of a slowly replaceable biological mediator.*

Multiple-Dose Administration—This refers to the administration of a succession of doses at intervals such that the drug does not leave the body completely in each interval between doses. The usual procedure in a multiple-dose regimen is to administer a drug repetitively with a constant dose interval, designated τ, with both dose and τ chosen so as to maintain the plasma concentration in the therapeutic

* Careful studies show that trace amounts of reserpine in the body outlast the effect and the duration of action may be related to these trace amounts. These residual amounts, however, are much smaller than are required to initiate the catecholamine-depleting action.

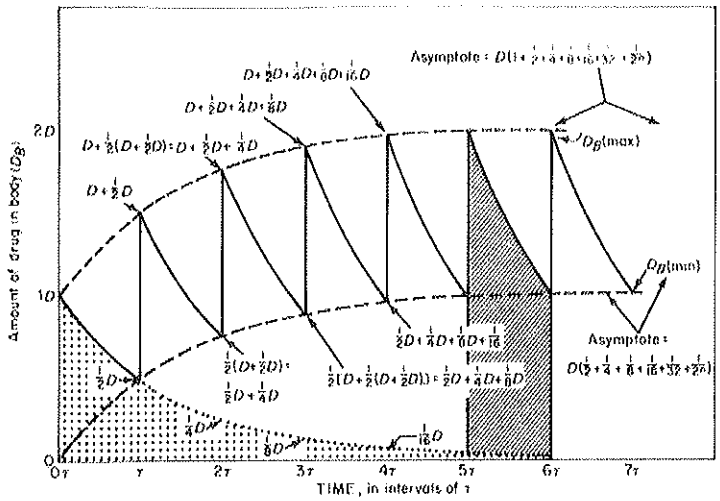


Fig 36-13. The accumulation of drug in the body during a regime of multiple dosing. Dose, D, is administered intravenously at intervals, τ, equal to the half-life, $t_{1/2}$. Thus, after each dose, the amount in the body, D_B, has decreased to half the previous peak amount at the time each dose is administered. When the cumulated amount in the body after injection reaches $2D$, the body content will fluctuate from $2D$ to $1D$ during each dose interval thereafter. Approximately 5 half-lives are required before this leveling off (plateau) of the body content occurs. The stippled area is the area under the elimination curve of a single injection, if no second dose had been given. The cross-hatched area is the area under the curve during a single-dose interval. The two areas are equal.

range. Some features of such repetitive dosing may be seen from the construction reproduced in Fig 36-13.

Accumulation and Plateau Principle—If the novice reader will make his own construction, it will aid greatly his understanding of the subject. In the construction, the amount of drug in the body, D_B, is plotted against time. Dose, D, is given repetitively, intravenously, at intervals such that $\tau = t_{1/2}$, in order to facilitate the construction. The first dose is given at $\tau = 0$; since it is given intravenously, the amount in the body rises to $D_B = 1$ essentially instantaneously. Immediately, D_B falls exponentially with the first-order kinetics of Eq 1, except that whole-body content, rather than C_p, is plotted. Since $\tau = t_{1/2}$, at τ, $D_B = \frac{1}{2}D$; when the next dose, D, is added, it brings the body content up to $D + \frac{1}{2}D$. During each dose interval, D_B falls exponentially to one-half the previous postinjection peak. As D_B rises after each administration, the rate (*not* the rate constant) of elimination rises proportionately, until eventually the amount eliminated during τ essentially equals the amount injected. The maximum and minimum values of D_B, $D_{B(max)}$ and $D_{B(min)}$, during τ, approach respective asymptotes, shown on the graph. As $t \to \infty$, $D_{B(max)} \to 2D$ and $D_{B(min)} \to D$. Thus, although D_B fluctuates between $D_{B(max)}$ and $D_{B(min)}$, once the asymptotes are approximated closely, D_B can be thought of as having reached a qualified *steady-state* condition, and the pharmacokinetics are sometimes called steady-state pharmacokinetics. Also, D_b is said to have reached a *plateau*. It is important to note that the rate at which the plateau is reached is at exactly the same rate at which drug is eliminated from the body after a single dose. Thus, the exponentially falling line for the elimination of D given at $\tau = 0$ (had no further doses been given) is the mirror image of the line connecting the sequential $D_{B(max)}$s. The principle that *when the rate of absorption is fast compared to the rate of elimination ($k_a > 5k_{el}$) the rate at which the multiple-dose steady state is approached is determined only by k_{el}*, and is known as the *plateau principle*. This is the fundamental feature of one-compartment multiple-dose kinetics. It obtains irrespective of the value of τ. However, the plateau concentrations do depend upon τ (see below).

In Fig 36-13, the drug was administered intravenously, so that no time-dependent absorption had to be considered. When absorption is involved, the $C_{p(max)}$ is not as high as with intravascular administration, but is blunted and occurs with a latency after administration that is determined by k_a/k_{el}, just as in single-dose administration. The appearance of the C_p-time curve with multiple-dose administration is shown in Fig 36-14. The value of C_p at any time during multiple-dose administration can be calculated according to Eq 30.

$$C_p = \frac{fk_a}{V_d(k_{el} - k_a)}\left[\left(D^* e^{-nk_{el}t} + D \cdot \frac{1 - e^{-nk_{el}\tau}}{1 - e^{-k_a}}\right) - \left(D^* e^{-nk_{el}t} + D \cdot \frac{1 - e^{-nk_{el}\tau}}{1 - e^{-k_{el}\tau}}\right)\right] \quad [\text{wt} \cdot \text{vol}^{-1}] \quad (30)$$

where n is the nth dose, τ is the dose-interval, t is the time since the last dose, D is the maintenance dose, D^* is the initial or loading dose (see below) and f is the fraction absorbed (bioavailability factor). With this equation, C_p, rather than D_B, is calculated; however, it will be recalled that

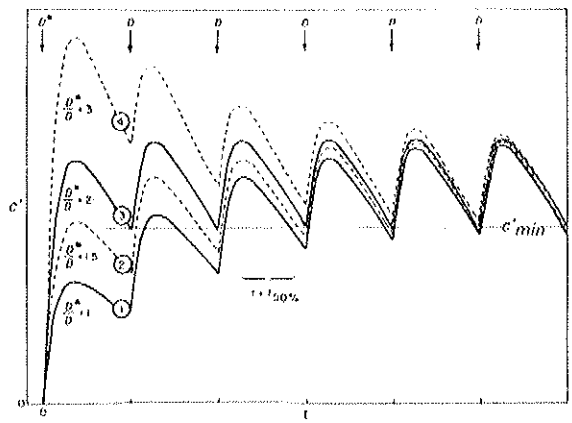


Fig 36-14. Time course of the plasma concentration of a drug administered according to a multiple-dose schedule. C' (ordinate): concentration; t (abscissa): time; D^*: initial dose; D: maintenance dose; τ: dose-interval (equal to $t_{1/2}$ in this illustration); C'_{min}: minimum concentration after each dose (same as $C_{p(min)}$ in text) (courtesy, Krüger-Thiemer[7]).

$C_p^0 = D/V_d$, and similarly, $C_p = D_B/V_d$, so that the equation easily is modified to calculate either C_p or D_B and the same principles apply in either form.

It is important to know how many half-lives must transpire before the plateau is approached closely enough to be considered complete for practical purposes. The value of $D_{B(min)}$ is approximately 93% complete at 4τ and 97% at 5τ; $D_{B(max)}$ is 97% at 4τ and 98.5% at 5τ. Thus, it may be stated that, for practical purposes, the plateau state is reached in approximately 5 half-lives, provided $k_a > 5k_{el}$. This is another form of the *plateau principle. The principle applies whenever the steady state conditions are perturbed;* that is, *5 half-times will be required to reach a new plateau,* whether the plasma concentration is rising or falling to a new plateau (see Fig 36-14).

Maximum and Minimum Concentrations—During multiple dosing, $C_{p(max)}$ and $C_{p(min)}$ are described by Eqs 31 and 32:

$$C_{p(max)n} = \frac{C_p^0\,(1 - e^{-nk_{el}t_a})}{1 - e^{-k_{el}\tau}} \qquad [\text{wt} \cdot \text{vol}^{-1}] \qquad (31)$$

$$C_{p(min)n} = \frac{C_p^0\,(1 - e^{-nk_{el}\tau})}{1 - e^{-k_{el}\tau}} \qquad [\text{wt} \cdot \text{vol}^{-1}] \qquad (32)$$

where n is the nth dose, C_p^0 is the concentration that would have occurred from instantaneous absorption and distribution (obtained by extrapolation of the elimination curve to zero time) and t_a is the absorption time. The term C_p^0 may be replaced by fD/V_d^{ss}. During the plateau state, $1 - e^{nk_{el}\tau}$ becomes $e^{-k_{el}\tau}$, and $C_{p(max)}$ and $C_{p(min)}$ are designated C_{max}^{ss} and C_{min}^{ss}, respectively. The equation is valid only when $k_a > 5k_{el}$. It can be seen that $C_{p(max)}$ is determined by both k_a and k_{el} (k_a shows itself only indirectly, in t_a) and $C_{p(min)}$ by k_{el}. The greatest difference between $C_{p(max)}$ and $C_{p(min)}$ occurs when the drug is given intravenously; when $\tau = t_{1/2}$, after intravenous injection, $C_{p(max)}/C_{p(min)}$ theoretically is equal to 2. With extravascular administration, the ratio is always less than that with intravenous administration, the ratio being determined by k_a/k_{el}. As k_a/k_{el} decreases, $C_{p(max)}/C_{p(min)}$ decreases.

Average Concentration and Body Content—The average concentration during the plateau state is described by Eq 33.

$$C_{p(ave)} = \frac{fD}{V_d k_{el}\tau} = \frac{1.44\,t_{1/2}fD}{V_d\tau} \qquad [\text{wt} \cdot \text{vol}^{-1}] \qquad (33)$$

The coefficient 1.44 is the reciprocal of 0.693 in Eq 3. The term $C_{p(ave)}$ is a time-averaged concentration and therefore is really a mean concentration. Since $C_p = D_B/V_d$, it follows that

$$D_{B(ave)} = \frac{fD}{k_{el}\tau} = \frac{1.44\,t_{1/2}fD}{\tau} \qquad [\text{wt}] \qquad (34)$$

It is self-evident that the plasma concentration, or amount of drug in the body, is directly proportional to the fraction of drug absorbed (f, bioavailability factor). The appearance of f in these equations and Eq 30, however, serves as a reminder that a change from one drug product to another with a different bioavailability, f', will be accompanied by changes in $C_{p(ave)}$ and $D_{B(ave)}$, as well as in the maxima and minima. The equations also reemphasize that a change in $t_{1/2}$ (or k_{el}) will affect $C_{p(ave)}$ and $D_{B(ave)}$, all other factors being held constant. Since k_{el} and f (and sometimes V_d in relation to weight) vary from patient to patient, the dosage of certain drugs always needs to be ascertained with laboratory assistance and acumen. The effects of changes in τ are discussed below.

Importance of Dose-Interval—The ratio $C_{p(max)}/C_{p(min)}$ depends on the dose-interval, τ. If the interval is increased and the dose is unchanged, $C_{p(max)}$, $C_{p(min)}$ and $C_{p(ave)}$ all decrease, but $C_{p(max)}/C_{p(min)}$ is increased. If τ is decreased, then $C_{p(max)}$, $C_{p(min)}$ and $C_{p(ave)}$ increase, but $C_{p(max)}/C_{p(min)}$ is decreased. This is shown in Fig 36-15. To avoid a change in $C_{p(ave)}$ consequent to a change in τ, the dose may be changed appropriately, in accordance with Eqs 32 and 33. Nevertheless, the wider fluctuations between $C_{p(max)}$ and $C_{p(min)}$, when τ is lengthened, cannot be avoided simply by adjusting the dose (see Fig 36-16, broken lines). If $C_{p(min)}$, rather than $C_{p(ave)}$, is held constant, the fluctuations become even larger (Fig 36-16, solid lines), and the hazard of the

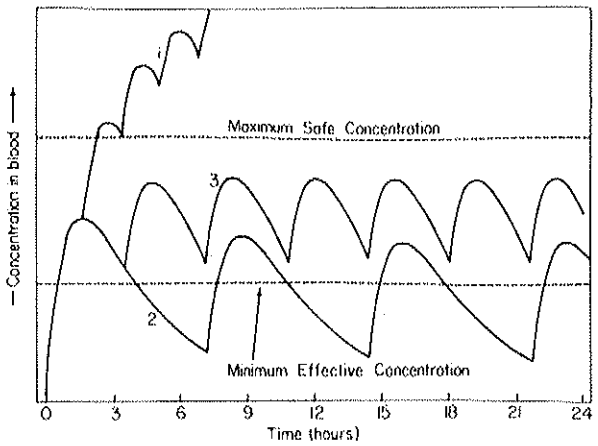


Fig 36-15. The effect of the dose interval on the time course of the plasma concentration of a drug administered in a multiple-dose regimen. $D^* = 4$, $D = 3$ and $k_a/k_{el} = 3$. The dose interval is 1.7 hr in Curve 1, 7.7 hr in Curve 2 and 3.8 hr in Curve 3 (courtesy, Notari[8]).

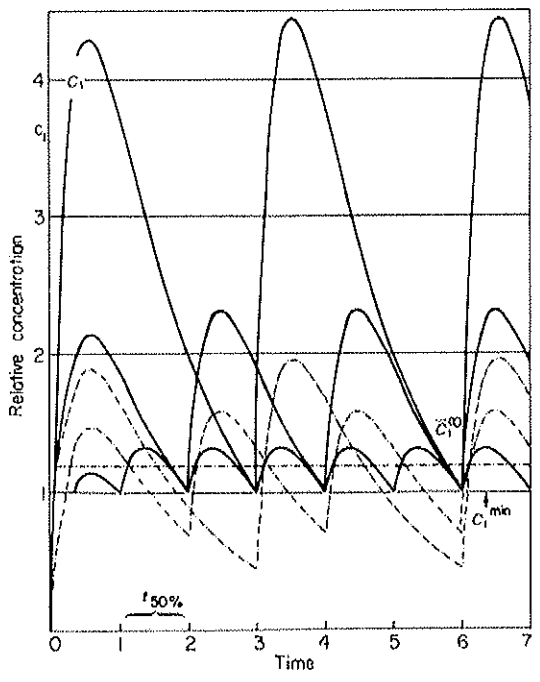


Fig 36-16. Fluctuations in the plasma concentration of a drug when the dose interval is changed but the dose is altered to maintain the same minimal (solid lines) or average (broken lines) concentration during maintenance. C_1^{min} is the minimal concentration (corresponding to $C_{p(min)}$ in the text) and $\bar{C}_1^{\infty}$ is the average concentration during maintenance (corresponding to $C_{p(ave)}$ in the text). Time is in multiples of the half-life (courtesy, Krüger-Thiemer,[7] adapted).

concentration reaching the toxic range is increased. Conversely, *the greater the number of divided doses, the smaller the fluctuations in plasma concentration.* For drugs with a narrow therapeutic range, it is usually inadvisable to dose at intervals longer than $t_{1/2}$. With digitoxin, τ is much smaller than $t_{1/2}$, and the fluctuations in plasma concentration are consequently less than 10%. However, for drugs with a high therapeutic index and which do not require a steady plasma concentration for an adequate therapeutic action, dose intervals much larger than $t_{1/2}$ may be used conveniently. Penicillin G is such a drug; it is more convenient to give large doses at 4-hr intervals, or longer, than at 30- to 60-min intervals ($t_{1/2}$ = 30 to 60 min).

Cumulation Ratio and Persistence Factor—From the above, it is evident that the drug cumulated in the body during the repetitive administration approaches different amounts (asymptotes) in the plateau state according to the magnitude of τ in relation to $t_{1/2}$ (or k_{el}). The dose-interval must be a convenient interval that not only is easy for the patient or medical and paramedical personnel to keep track of but also one which does not subject the patient to an annoying or difficult number of doses per day. Furthermore, $t_{1/2}$ varies from patient to patient. Consequently, it is rare when $\tau = t_{1/2}$, although it is sometimes close enough that the difference is inconsequential. Therefore, it is important to be able to estimate the extent of cumulation with any dose interval in any patient. This can be done with information derived from a single dose, by means of the *accumulation factor*, r_a.

$$r_a = \frac{1}{1 - e^{-k_{el}\tau}} \qquad \text{[no units]} \qquad (35)$$

The component factor, $e^{-k_{el}\tau}$, is the *persistence factor*, r, which is the fraction by which C_p or D_B falls during the dose interval. When the plateau, or steady state, is reached the cumulated plasma concentration or body content will be larger than that from the first dose by a factor known as the *cumulation ratio* (or *drug amount ratio*), R_c.

$$R_c = \frac{1}{k_{el}\tau} = \frac{1.44\, t_{1/2}}{\tau} = \frac{\bar{C}_\tau^{ss}}{\bar{C}_0^\infty} \left(\text{or} \frac{\bar{D}_{B\tau}^{ss}}{\bar{D}_{B0}^\infty} \right) \qquad \text{[no units]} \qquad (36)$$

where $\bar{C}_\tau^{ss}$ is the mean concentration during one dosage interval during the steady state and $\bar{C}_0^\infty$ is the mean concentration from $t = 0$ to $t = \infty$ after a single dose; $\bar{D}_{B\tau}^{ss}$ and $\bar{D}_{B0}^\infty$ are the corresponding respective body contents. Since both $\bar{C}_0^\infty$ and $\bar{C}_\tau^{ss}$ can be estimated from the AUC, it is appropriate to discuss this further.

Area under Curve (AUC)—The area under the monoexponentially falling, single-dose plasma concentration-time curve is the integral of the differential form of Eq 1, from $t = 0$ to $t = \infty$:

$$AUC^{0\to\infty} = \bar{C}^{0\to\infty} = \int_0^\infty C dt = \int_0^\infty C_p^0 e^{-k_{el}t} dt = \frac{C_p^0}{k_{el}} \qquad \text{[wt} \cdot \text{vol}^{-1} \cdot \text{time]} \qquad (37)$$

Although the units are concentration times time, the value is equal to the time-averaged concentration and hence is called the *average concentration* $\bar{C}_0^\infty$, although it is more appropriately a *log-mean* concentration. If the amount of drug in the body is used, instead of plasma concentration, the AUC is equal to the *time-averaged body content*. The average body content could, of course, be calculated from $\bar{C}_0^\infty$ by multiplying by V_d.

Even when two or more exponential processes act additively on the plasma concentration (or body content), as in absorption plus elimination, the $AUC^{0\to\infty}$ equals $\bar{C}_p^\infty$ (or $\bar{D}_B^\infty$). The interested student may verify this by integrating any of Eqs 28–30. In the two-compartment system (see below), $AUC^{0\to\infty}$ for a plasma concentration-time curve correctly equals $\bar{C}_p^\infty$; however, $\bar{D}_B^\infty$ cannot be calculated from $\bar{C}_p^\infty$, because the plasma concentration differs from the average body concentration.

Since $AUC^{0\to\infty} = C_p^0/k_{el}$ in the one-compartment system, it is obvious that AUC does not provide any new information that otherwise cannot be obtained, as by back-extrapolation or regression analysis. Nevertheless, AUC frequently is used in lieu of C_p^0/k_{el}. For example, in the determination of the bioavailability factor, f, the $AUC^{0\to\infty}$, after *extra*vascular administration ($AUC_{ev}^{0\to\infty}$), divided by the AUC after *intra*vascular administration ($AUC_{iv}^{0\to\infty}$) is equal to f.

The term $AUC^{0\to\infty}$ is not the only AUC that may be used in pharmacokinetics. The AUC during different time intervals, under supposedly steady-state conditions, could be employed to detect time- or concentration-related changes in clearance (eg, see Eqs 11 and 27). During the plateau, or steady state, the AUC during one dose interval (AUC^{ss}) is of special interest. To evaluate $AUC^{0\to\infty}$ requires many samples taken over a long period of time, which is an inconvenience to the subject or patient. The value of AUC^{ss} can provide the same derived information with fewer samples and less time. This is because $AUC^{ss} = AUC^{0\to\infty}$. Thus, in Fig 36-13, the stippled area, which is $AUC^{0\to\infty}$, would be equal to the cross-hatched area, AUC^{ss}, except for the negligible stippled area that remains after 5τ. At $t = \infty$, the two areas would be essentially identical. In this comparison of AUCs, the identical areas do not mean that $\bar{C}_p^\infty$ is identical to $\bar{C}_p^{ss}$, but it does enable AUC^{ss} to be used to calculate values of single-dose parameters and *vice versa*.

Constant Infusion and Sustained Release—A constant infusion or sustained release of a drug may be regarded as a series of minidoses given at infinitely short dose intervals. When infusion is intravascular, the plasma concentration will rise in logarithmic fashion with the same time course and cumulation factor as with multiple dosing, ie, with a rate constant of k_{el}. Thus, the plateau principle applies equally to constant infusion and multiple dosing. After discontinuation of infusion, the plasma concentration falls exponentially with a rate constant k_{el}, in accordance with Eq 1. These principles are illustrated in Fig 36-17.

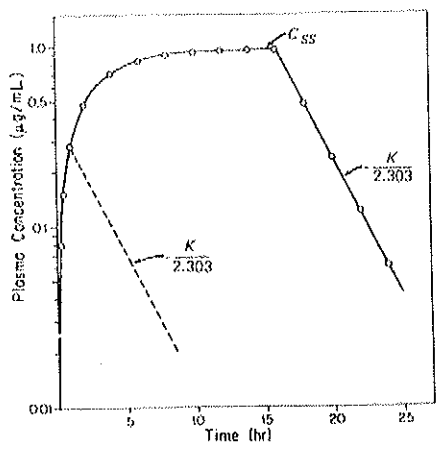


Fig 36-17. Semilogarithmic plot of plasma concentration during and after cessation of a constant intravenous infusion of a drug in a one-compartment system. Whether infusion is stopped prior to the attainment of a plateau or after, the plasma concentration will fall log-linearly with a slope of $-0.434k_{el}$. In the figure, K is k_{el} and $1/2.303 = 0.434$. C_{ss} is the steady-state concentration, C_p^{ss} (courtesy, Gibaldi and Perrier[9]).

The steady-state plasma concentration, C_p^{ss}, is equal to the infusion rate divided by the whole body clearance:

$$C_p^{ss} = \frac{R^0}{Cl_{tot}} = \frac{R^0}{V_d k_{el}} \qquad [\text{wt} \cdot \text{vol}^{-1}] \qquad (38)$$

where R^0 is the infusion rate. The term V_d must be expressed in the same volume units as R^0; Cl_{tot} and R^0 must be in the same time units as k_{el}.

With sustained-release dosage forms, in which the release is approximately constant for long periods of time, the pharmacokinetics are like those of constant infusion.

Loading and Maintenance—In Fig 36-13, $D_{B(max)} \rightarrow 2D$; consequently, had $2D$ been given for the first dose and D thereafter, the plateau condition would have been reached immediately. This illustrates the *principle of loading*. The same effect of loading is shown by curve 3 in Fig 36-14; in both these figures, $\tau = t_{1/2}$. The initial dose is called the *loading dose*, D^*, and each subsequent dose is called the *maintenance dose*, D. Since it takes about 5 half-lives to reach the plateau state, it is very important to use a loading dose with drugs that have long half-lives or in situations in which it is desirable that the optimal therapeutic concentration be reached rapidly.

The loading dose, D^*, should approximate the amount of drug in the body which will be contained during maintenance (ie, the plateau state). The most direct way to calculate D^* is with the equation

$$D^* = \frac{V_d \cdot C_{p(max)}^{ss}}{f} \qquad [\text{wt}] \qquad (39)$$

assuming that V_d^{ss} and $C_{p(max)}^{ss}$ are both known. A first dose so calculated achieves a $C_{p(max)}$ that is equal to that at the steady state only for *intra*vascular administration. After *extra*vascular administration $C_{p(max)}$ is less than that after *intra*vascular administration and hence the loading dose is proportionately smaller. With some intravascularly administered drugs, the loading dose is calculated deliberately to be less than that calculated by Eq 39. Among reasons for choosing a lower dose than that calculated by Eq 39 is that the effects of the first of a series of doses often elicits greater responses than do subsequent doses, because reflex, hormonal and other counter-regulatory effects have not had enough time to come into full play. This practice applies even to some *extra*vascularly administered drugs, such as prazosin. Consequently, C_p^{ss}, or even $C_{p(min)}^{ss}$, may be used *in lieu* of $C_{p(max)}^{ss}$. It must be remembered that with such underloading the steady state is not achieved fully with the loading dose. With drugs which have a very low and erratic therapeutic index and potentially fatal toxicity, the loading dose may be divided into smaller doses, to be given at various intervals before the first maintenance dose; this permits monitoring of both C_p and clinical effects during loading and allows an assessment of whether the intended maintenance dose is correct. Fractional loading also is used when a drug with a low therapeutic index has a significant distribution phase, such that toxic plasma concentrations occur before distribution equilibrium occurs. With some drugs, an appropriate V_d^{ss} is not known, thus making Eq 39 inapplicable. With such drugs, D^* can be calculated from traditional, empirical maintenance doses by means of the equation

$$D^* = \frac{D}{(1 - e^{-k_a \tau})(1 - e^{-k_{el}\tau})} \qquad [\text{wt}] \qquad (40)$$

The equation correctly applies only when $k_a > 3k_{el}$. Also, D^* can be calculated according to

$$D^* = fD/R_c = 1.44 f D t_{1/2}/\tau \qquad [\text{wt}] \qquad (41)$$

where R_c is the cumulation ratio (see Eq 36).

The time course of the plasma concentration after different loading doses is shown in Fig 36-14. When $D^* = 2D$, the plateau maintenance concentration is approximated closely when $\tau = t_{1/2}$ but is smaller than 2 when $\tau < t_{1/2}$ and greater when $\tau > t_{1/2}$.

In Fig 36-14, it should be noted that if the loading dose is not optimal, either too low or too high, the plateau state is approached with the same time course as when no loading dose is given.

When a constant intravenous infusion is used, the principle of loading also applies, because the plateau principle applies; loading may be accomplished with one or more rapid intravenous doses, called boluses or slugs, or by an initial period of rapid infusion to bring the plasma concentration to the maintenance level. The loading dose can be calculated from Eq 39 or the infusion rate and half-time, as

$$D_0^* = \frac{R_0 t_{1/2}}{0.434 \log 2} \qquad [\text{wt}] \qquad (42)$$

Open Two-Compartment Model

The one-compartment model adequately describes the pharmacokinetics of many drugs. However, with an even larger number of drugs, after intravenous administration, the decline in plasma concentration is not monoexponential but rather manifests two or more monoexponential components which are discernible in the semilogarithmic plot of C_p versus time. The most common is a decline which manifests two components; the open two-compartment model most adequately describes such pharmacokinetics. Other models having more compartments or other complexities will be mentioned later briefly.

Description of the Model—In the open two-compartment model, the body is considered to comprise two compartments in dynamic equilibrium, as depicted in Fig 36-18. The compartment into which the drug is directly absorbed and from which the drug is eliminated is called compartment 1, or the *central compartment*. The blood is a part of this compartment, is the transporting and distributing medium and is the medium actually sampled for chemical and pharmacokinetic analysis; consequently, compartment 1 is some-

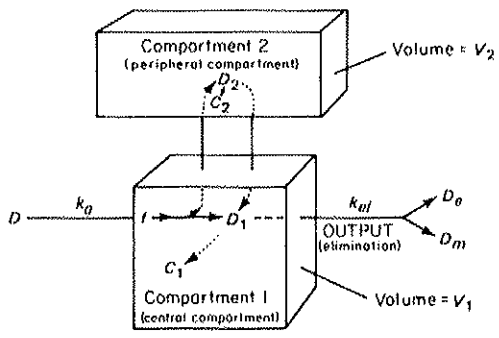


Fig 36-18. Diagram of open two-compartment pharmacokinetic model. An amount of drug, fD, is absorbed from the administered dose, D, with a first-order rate constant of k_a into compartment 1 of volume V_1. Some of the absorbed drug enters compartment 2 with a first-order rate constant of k_{12} and is returned into compartment 1 with a first-order rate constant of k_{21}. D_1 is the amount of drug in compartment 1 and D_2 in compartment 2; C_1 and C_2 are the respective concentrations in compartments 1 and 2 ($C_1 = C_p$). Drug is eliminated from compartment 1 with a first-order rate constant, k_{el}, which, however, is obscured by the lag in transfer of drug from compartment 2 to compartment 1. D_e is the amount excreted into urine, feces, expired air, sweat, milk, etc; D_m is the amount of drug metabolized. The relative volumes of V_1 and V_2 may vary greatly, V_1 sometimes being the larger and other times the smaller.

times misleadingly called the blood or plasma compartment, even though the erythrocytes or plasma proteins may sometimes behave kinetically as though they were part of compartment 2. In the simple two-compartment model, compartment 2 is closed and communicates with the environment only through the central compartment, being, as it were, peripheral to the events of absorption and elimination; consequently, it is called the *peripheral compartment*. Sometimes, it also is called the tissue compartment, which is misleading, since usually some tissues, or certain cell types within otherwise peripheral tissues, may be kinetically in compartment 1. It is important to reiterate that the compartments are fictive and are defined by the kinetic behavior of the drug within the body and not necessarily by identifiable anatomical entities. To avoid confusion and to enable a simple numerical designation of model components and distribution rate constants by number, the terms compartment 1 and compartment 2 will be used hereafter.

The movement of drug between compartments is defined by characteristic first-order rate constants. The subscript indicates the direction of movement; thus k_{12} (subscript one-two, not twelve) indicates movement from compartment 1 to compartment 2 and k_{21} the reverse direction. The constants k_a and k_{el} are entirely analogous to the like-designated respective absorption and elimination rate constants of the one-compartment model. However, k_{el} is not observed directly from the decline in plasma concentrations, since both the characteristic overall rate of the elimination processes and the rates of diffusion into, and recruitment from, compartment 2 combine to control the rate of decline in plasma concentration (see below). Once an infinitesimal amount of drug is absorbed, all processes occur simultaneously, ie, in parallel. Nevertheless, since the various processes have different time constants, one process will run its course to a practical end earlier than another, and events may be thought of as occurring sequentially, with overlap, in the order; absorption, distribution and elimination. So long as $k_a > (k_{12} + k_{21})/k_{21} > k_{el}$, the terminal phase will be a steady decline in concentration (see Fig 36-19), during which the distribution ratio, C_1/C_2, will be constant.

Absorption—Absorption does not differ from that in the open one-compartment model and does not require further description. However, the determination of absorption characteristics from the log plasma concentration-time curves is complicated by the distribution phase, and the method of residuals (page 733) entails the resolution of three, rather than two, components (see below).

Distribution and Elimination—After the intravascular administration of a drug which obeys two-compartment kinetics, the plasma concentration falls in a complex two-process fashion, but in an arithmetic plot the two components may not always be evident to the eye. When concentration-time data are plotted semilogarithmically, however, the separate processes of distribution and elimination are identified easily by the method of residuals (back-feathering, page 733 and Fig 36-8), if the rate of distribution exceeds significantly that of elimination. In Fig 36-19, such a resolution has been made for the drug pralidoxime. In the figure, it may be seen that after 2 hr the curve assumes a log-linear character. The assumption is made that the distribution phase essentially is complete and a pseudoequilibrium has been reached between the two compartments. Therefore, the late log-linear segment of the line, with the slope -0.434β, represents the elimination phase. If this line is subtracted from the nonlog-linear portion of curve, the distribution phase is the residual line. In order to do this, the log-linear segment is back-extrapolated. From this extrapolated line are obtained the antilogs to be subtracted from the temporally corresponding antilogs on the unresolved, original curve. The respective differences, or residuals,

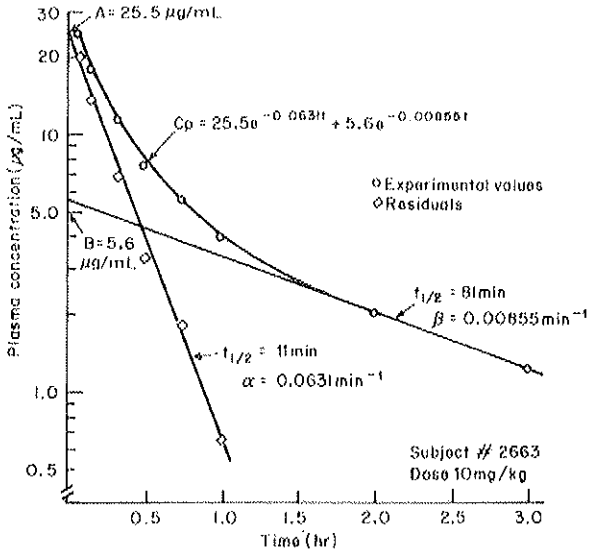


Fig 36-19. Resolution of the plasma concentration curve for pralidoxime into its distribution and elimination components after intravenous administration. Note that plasma concentration is plotted on a logarithmic scale. The time constant for the elimination phase is determined from the slope, -0.434β; it is a hybrid constant and β is not the same as k_{el} (see text). Likewise, the time constant for distribution, α, is obtained from the slope, -0.434α, of the distribution line; α is also a hybrid constant (courtesy, Gibaldi and Perrier[9]).

then are plotted semilogarithmically to reveal the log-linear line that represents distribution only. From the log-linear properties of the separate, but algebraically additive, lines representing the two processes of distribution and elimination, it may be inferred that the equation for the original compound curve was

$$C_1 = Ae^{-\alpha t} + Be^{-\beta t} \qquad [\text{wt} \cdot \text{vol}^{-1}] \tag{43}$$

where C_1 is the concentration of drug in compartment 1 (the central compartment), α and β are first-order rate constants for the distribution and elimination phases, respectively and A and B are fictive plasma concentrations to be discussed on page 740. The constant β describes the late rate of disappearance of drug from compartment 1 but is not the same as k_{el} (see below). It is the rate constant from which the biological half-life is calculated in a two-compartment system ($t_{1/2} = 0.693/\beta$).

Hybrid and Prime Kinetic Parameters—In Fig 36-19, the slope of the late, slower elimination line is -0.434β, where β is a first-order time constant for elimination. However, β is determined not only by the rate capacities of the irreversible elimination processes but also by the rates at which drug is transferred out of and back into compartment 1. Therefore, β is a compound, or hybrid, rate constant. It is equal to the fraction of drug in the central compartment, sometimes designated as f^*, in the postdistributive (elimination) phase times the elimination constant, k_{el}, for the central compartment. Thus

$$\beta = f^* k_{el} \qquad [\text{time}^{-1}] \tag{44}$$

Alpha, α, is a hybrid constant that combines k_{21}, k_{el} and β:

$$\alpha = \frac{k_{21} k_{el}}{\beta} \qquad [\text{time}^{-1}] \tag{45}$$

Interestingly, the equation for α does not include k_{12}, although f^* does depend upon $(k_{12} + k_{21})/k_{21}$. The sum of α and β can be expressed entirely in terms of prime constants:

$$\alpha + \beta = k_{12} + k_{21} + k_{el} \qquad [\text{time}^{-1}] \tag{46}$$

However, these prime constants cannot be determined directly and must be derived from the hybrid constants that are obtainable from graphical or regression analysis. The formulae are

$$k_{el} = \frac{A + B}{\frac{A}{\alpha} + \frac{B}{\beta}} \qquad [\text{time}^{-1}] \tag{47}$$

$$k_{12} = \frac{AB\,(\beta - \alpha)^2}{(A + B)(A\beta + B\alpha)} \qquad [\text{time}^{-1}], \tag{48}$$

and

$$k_{21} = \frac{A\beta + B\alpha}{A + B} \qquad [\text{time}^{-1}] \tag{49}$$

where A and B are the zero-time intercepts of the residual distribution line and the postdistributive (elimination) line, respectively. Each represents a fictive concentration that describes a limit when the other variable is set to zero (ie, the other process is nonexistent).

The volume of compartment 1 (central compartment) can be obtained from C_p^0 (ie, $V_1 = fD/C_p^0$). From the fictive concentration, B, the apparent volume of distribution during the postdistributive phase can be calculated, since $A + B = C_p^0$. From $A - B$ may be obtained the value of compartment 2. (Volumes of distribution are discussed below.) C_p^0 can be determined more accurately by summing the two log-linear extrapolates than from extrapolation of the unresolved curve. The coefficients A and B are also hybrid, since the value of B depends upon all of k_{21}, k_{12} and k_{el}.

Volumes of Distribution—The volume of distribution, V_d, of a drug is a useful pharmacokinetic parameter that relates C_p to D_B (see page 727). Even though it is fictive, it provides not only some insight into distribution but also importantly relates to the rate of clearance of drug from plasma, and changes in pathological conditions reveal changes in the physiological-biochemical conditions. By means of the distribution coefficient, Δ', data from one patient may be applied to others of different body weights (see page 728).

In the open two-compartment system, the determination of V_d is complicated by the slow attainment of distribution "equilibrium" (ie, steady state) between two compartments, and the volume of distribution is changing continually during the distribution phase. It is especially important to know V_d during the postdistribution phase (in which case V_d only applies during postdistribution times) or to estimate V_d by methods that cancel the distributive factors.

Theoretically, the most accurate method for estimating V_d is known as the *steady-state* method, of which there are three variations. In this, the ideal procedure is to give a continuous intravenous infusion until the steady state (ie, plateau) is reached. During the steady state, the amount of drug in the peripheral compartment (compartment 2) is constant. Under these conditions

$$V_d^{ss} = \frac{k_{12} + k_{21}}{k_{21}} \cdot V_1 \qquad [\text{vol}] \tag{50}$$

Note that V_d^{ss} is independent of k_{el} and β. There are, however, several disadvantages to this approach, the principal ones being that for most drugs the steady state is reached only after prolonged infusion, since 5 or more half-lives often will require days of infusion, and that V_1, k_{12}, k_{21} and β need to be determined. This can be done by discontinuing infusion and resolving the curve of the declining plasma concentration into its component parts. Fortunately, the same information can be obtained from the mean plasma concentration during one dose-interval at steady state, C^{ss}. In this,

$$V_d^{ss} = \frac{fD(k_{12} + k_{21})}{C^{ss}k_{21}k_{el}\tau} \qquad [\text{vol}] \tag{51}$$

where k_{el} is the rate of elimination from the central compartment. Provided that elimination occurs only from the central compartment, Eqs 50 and 51 are valid for any n-compartment model. This method has the same disadvantage as the infusion method in that dosing must be continued to the steady state, which, however, with repetitive dosing is more comfortable and less expensive than continuous infusion. An advantage is that extravascular routes may be employed and that only one dose-interval need be sampled, thus making the determination of V_d^{ss} applicable to drugs with long half-lives.

The value of V_d^{ss} also can be determined from areas under the curve (AUC) during and after constant intravenous infusion

$$V_d^{ss} = \frac{D_\Sigma \cdot AUC_{t(ss)}}{C^{ss} \cdot AUC^{0\text{-}\infty}} \qquad [\text{vol}] \tag{52}$$

where $t(ss)$ is the time to reach steady state, D_Σ is the cumulated dose at $t(ss)$, $AUC_{t(ss)}$ is the area under the plasma concentration-time curve from $t = 0$ to $t = t(ss)$ and $AUC^{0\text{-}\infty}$ is the total area under the curve from $t = 0$ to $t = \infty$, providing that the infusion is stopped at the achievement of steady state or that the AUC, during any overrun into the plateau state, is eliminated from the determination of $AUC^{0\text{-}\infty}$. The method has the advantage that the determination of k_{12}, k_{21}, k_{el} or V_1 is not necessary.

A second method of determining V_d is that in which V_d is calculated from V_1, k_{el} and β:

$$V_{d(\beta)} = \frac{V_1 k_{el}}{\beta} \qquad [\text{vol}] \tag{53}$$

The designation $V_{d(\beta)}$ indicates the method of calculation. The rationale for the method is the valid assumption that plasma and tissue concentrations decline in parallel during the postdistributive phase, so that the distribution ratio, which will be equal to Δ', is constant after the distributive phase has come to completion. The method has been shown to yield the same values for V_d as one based on area:

$$V_{d(area)} = \frac{fD}{AUC^{0\text{-}\infty}} = \frac{fD}{(A/\alpha + B/\beta)\beta} = V_{d(\beta)} \quad [\text{vol}] \tag{54}$$

The method is independent of the route of administration, so long as the fraction absorbed, f, is used.

On page 739, on which the parameters derived from curves such as that in Fig 36-19 were discussed, it was pointed out that the zero-time extrapolates A and B were fictive concentrations from which apparent volumes of distribution could be obtained. The extrapolate β gives a volume known as $V_{d(extrap)}$:

$$V_{d(extrap)} = \frac{D}{B} \qquad [\text{vol}] \tag{55}$$

The method does not take into account the effect of process k_{21} to limit the size of the peripheral compartment and hence tends to overestimate D_B, except at zero time. However, it has the advantage of rapid determination.

The value of $V_{d(area)}$ is the most correct approximation of V_d to apply to the postdistribution phase and $V_{d(ss)}$ is correct for constant infusion at steady state but otherwise underestimates D_B. By magnitude, these three volumes of distribution rank as follows: $V_d^{area} > V_d^{ss} > V_1$.

Clearance—The definition and concept of clearance can be found on page 729. The definition of clearance applies

whether the elimination occurs in a one- or multi-compartment system, hence clearance is model-independent. However, mathematical identities of clearance do depend on the model. In the open two-compartment model, β and $V_{d(area)}$ are applicable in the calculation of total body clearance:

$$Cl_{tot} = \beta V_{d(area)} \qquad [\text{usually mL} \cdot \text{min}^{-1}] \qquad (56)$$

Since it is customary to express clearance in units of mL/min, β must be expressed in min and $V_{d(area)}$ in mL. An analogous formula is based on the condition of the model that elimination occurs only from the central compartment, so that the applicable volume and elimination-rate constant are used:

$$Cl_{tot} = k_{el} V_1 \qquad [\text{mL} \cdot \text{min}^{-1}] \qquad (57)$$

Cl_{tot} also can be expressed in terms of α, A, β, B and D:

$$Cl_{tot} = \frac{D}{A/\alpha + B/\beta} \qquad [\text{mL} \cdot \text{min}^{-1}] \qquad (58)$$

Absorption Plus Distribution and Elimination—After extravascular administration in a two-compartment system, there are three first-order processes occurring simultaneously: absorption, distribution and elimination. These processes all add algebraically, as follows

$$C_p = Ae^{-\alpha t} + Be^{-\beta t} - C_p^0 e^{-k_a t} \qquad [\text{wt} \cdot \text{vol}^{-1}] \qquad (59)$$

They can be resolved by various methods, of which the easiest is the method of residuals already illustrated in Figs 36-8 and 36-19. However, in a two-compartment system, the first residual line is a compound line (absorption + distribution) and must be resolved further into its two component lines. Figure 36-20 is an example of the method of residuals applied to two-compartment data. The first step is the subtraction of the late postdistribution (elimination) line (with slope -0.434β) from the curve, which leaves a two-component residual curve. This residual curve has a late, postabsorptive log-linear segment of slope -0.434α. If the absorption segment of the curve of residuals is subtracted from the extrapolated α-line, a log-linear second residual line with a slope of $-0.434k_a$ will be generated. The extrapolated intercepts A and B have the meanings previously discussed. The zero-time intercept of the absorption residual line is equal to C_p^0 and hence, theoretically equals $A + B$. Kinetic parameters other than α, A, β and B are calculated by means of Eqs 44 and 45. The absorption parameters for other routes of absorption can be determined similarly, except with certain sustained-release dosage forms, which release approximately at a steady rate over long periods of time.

In the example illustrated by Fig 36-20, only two or three points each could be used for establishing the log-linear segments of the residual distribution and absorption lines, which, therefore, may be in considerable error. This indicates the importance of taking frequent enough samples, especially during the absorption and distribution phases, to provide reliable kinetic data.

Multiple-Dose Administration—Equations 30–34, which describe various aspects of the fluctuating plasma concentrations in the one-compartment system, are complex. It may be appreciated that the additional complexities conferred by two compartments renders the analogous equations intricate and difficult to follow for the nonspecialist. However, one-compartment equations modified in minor ways apply to two-compartment systems with reasonable accuracy, when the distribution phase after one dose is approximately complete before the next dose is administered. Under these conditions, β may be substituted for k_{el} and $V_{d(area)}$ for V_d, to adapt one-compartment equations to two-compartment systems for rough approximations of the

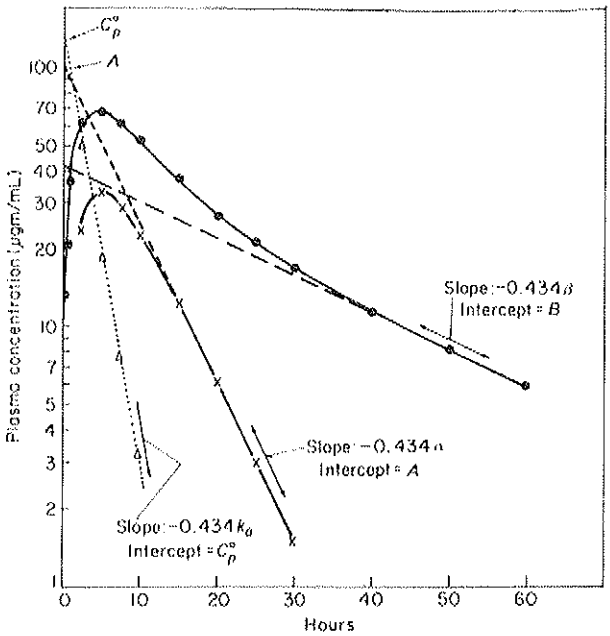


Fig 36-20. Resolution of absorption, distribution and elimination components of a concentration-time curve of a drug with two-compartment kinetics. The solid curve is a semilogarithmic plot of plasma concentrations. The method of residuals was used to resolve the component lines. The postdistribution, or elimination, line of slope -0.434β (— —) was subtracted from the concentration-time curve. The difference, or residual line (X—X) retained the absorption and distribution components. The log-linear segment of this line represents the postabsorption ("distribution") line, of slope -0.434α. A second residual line representing the absorption phase was obtained by subtracting the absorptive segment (first four points) of the first residual curve (X—X) from the extrapolated α line of slope -0.434α (- - - -) to give the residual absorption line of slope $-0.434k_a$ (· · · · ·). The zero-time intercepts of the extrapolated lines defined by k_a (· · · · ·), α (- - - -) and β (— —) are C_p^0, A and B, respectively (courtesy, data, Gibaldi and Perrier[9]).

two-compartment parameters and plasma concentrations. Thus,

$$C^{ss} = \frac{fD}{\beta V_{d(area)}\tau} \qquad [\text{wt} \cdot \text{vol}^{-1}] \qquad (60)$$

Adaptation of one-compartment equations for accumulation ratios and loading dose also usually gives values that satisfactorily approximate those calculated with more rigorous equations. The respective adapted equations are

$$R_c = \frac{1}{1 - e^{-\beta\tau}} \qquad [\text{no units}] \qquad (61)$$

and

$$D_0^* = R_c D \qquad [\text{wt}] \qquad (62)$$

where R_c, D_0^*, and D are the accumulation ratio, optimal loading dose and maintenance dose, respectively. In some instances, eg, when a rapid response to lidocaine is desired, a loading dose calculated with Eq 61 will be too low to provide adequate antidysrhythmic effects of the drug during the distribution phase. In this case, a loading dose can be approximated by use of the formula

$$D^* = V_1 C_p \qquad [\text{wt}] \qquad (63)$$

where D^* is the loading dose, V_1 is the volume of the central

compartment and C_p is the target (immediate central compartment) plasma concentration.

The rate at which the steady state is attained depends almost entirely on β. The plateau principle essentially applies, and approximately 5 half-lives, based on β, are required to reach the *steady state*. Essentially all precepts emanating from the one-compartment plateau principle are applicable if two-compartment β is used in place of one-compartment k_{el}.

Nonconformities and Miscellany

Fallibility of Assumptions—General pharmacokinetic concepts are applicable to many drugs without significant modification. Implicit in these concepts are certain assumptions which, however, do not apply to all drugs or drug recipients. Some of the basic assumptions are (1) the pharmacological effect is elicited by the drug administered (and which is being assayed in the blood), (2) the pharmacokinetic parameters remain constant with both time and dose and (3) the peak effect occurs when the concentration is at its peak at the site of action, binding and sequestration follow first-order kinetics and, in short, the models chosen for kinetic analysis are correct. When these assumptions are not valid, significant clinical consequences accrue, and theoretical and/or empirical modification of the models may be necessary. Therefore, it is worthwhile to examine some departures from the more common or commonly assumed behavior and some miscellaneous pharmacokinetic considerations not stressed elsewhere in this chapter.

Active Metabolites and Latentiation—Some drugs are biotransformed to a metabolite that has a pharmacological action like that of the parent drug. With these, the pharmacokinetics of each of parent drug and its metabolite may or may not be simple and easy to define, but the combined pharmacodynamic (and sometimes pharmacokinetic) action may rise and fall in a complex way because of the different time courses, distributions and routes of elimination of the two active molecules. For example, the anticonvulsant trimethadione (TMO) is un-ionized at body pH, is little excreted and has a V_d of about 600 mL/kg and a half-life of about 4 hr, whereas its anticonvulsant metabolite, dimethadione, is a weak acid, is excreted and excretion is affected by urine pH, has a V_d of 400 mL/kg and has a half-life of about 10 days. It is obvious that a study of the pharmacokinetics of TMO alone would be of little value in predicting a therapeutic regimen and precautions.

Two or more active metabolites may increase the complexity greatly. There are a few drugs in which it is only the metabolite, not the parent drug, that is active; with these, the relationship of pharmacokinetics to pharmacodynamics is simpler, provided that it is the metabolite that is followed. It is sometimes deliberately the practice to prepare a drug that is inactive with the intention that the drug be converted to an active metabolite once it is in the tissues. This practice is known as *latentiation*. Latentiation may be used when it is desired to slow down the rate of delivery of drug to the tissues, a kind of systemic sustained release, as it were, or when the active metabolite is locally toxic at the site of administration. Some drugs which generate active metabolites are shown in Table III. Not shown are drugs whose metabolites have no therapeutic activity but which have toxic or other pharmacodynamic activity.

The amount of a metabolite of a drug in the body at any one time depends upon both the rate of transformation of the drug to metabolite and the rate of disposition of the metabolite. The body content of metabolite will continue to rise so long as the content of precursor is high enough that the rate of biotransformation to metabolite exceeds the rate of elimination of the metabolite. When the concentration of drug or precursor falls to a level below which there is no longer a net gain in content of metabolite, the metabolite concentration will fall.

Table III—Some Drugs with Pharmacologically Active Metabolites

Parent Drug	Active Metabolite(s)
Acetohexamide	Hydroxyhexamide
Allopurinol	Alloxanthine
Aldophosphoramide	Phosphoramide mustard
Amitriptyline	Nortriptyline
Chloral Hydrate	Trichloroethanol
Chlordiazepoxide	Desmethylchlordiazepoxide, Demoxepam
Codeine	Morphine
Dacarbazine	5-Aminoamidazole-4-carboxamide
Diazepam	Desmethyldiazepam
Digitoxin	Digoxin
Flurazepam	Desalkylflurazepam
Fluorouracil	Fluorodeoxyuridine phosphate
Glutethimide	4-Hydroxyglutethimide
Imipramine	Desipramine
Lidocaine	Glycinexylidide
Meperidine	Normeperidine
Mephobarbital	Phenobarbital
Methyldopa	α-Methylepinephrine, α-methylnorepinephrine
Methamphetamine	Amphetamine
Phenacetin	Acetaminophen
Phenylbutazone	Oxyphenbutazone
Prednisone	Prednisolone
Primidone	Phenobarbital
Propoxyphene	Norpropoxyphene
Procainamide	*N*-Acetylprocainamide
Propranolol	4-Hydroxypropranolol
Spironolactone	Canrenone, Canrenoate
Sulfasalazine	Sulfapyridine
Tamoxiphen	4-Hydroxytamoxiphen
Trimethadione (TMO)	Dimethadione (DMO)

The kinetics of the fall in concentration depends upon which rate is faster, the elimination of drug precursor or the elimination of metabolite. If that of the drug is faster, the content of metabolite will rise above that of the drug, and the drug will soon disappear. This eventually leaves the content of cumulated metabolite to decline according to the kinetics of its own disposition.

In Fig 36-21, drug B illustrates the rate-limiting effect of the disposition of a metabolite. When the rate constant for the elimination of the drug or precursor is slower than that of the metabolite, as with drug A in Fig 36-21, the content of metabolite never reaches that of the drug and it eventually declines according to the kinetics of biotransformation of the drug. That is, the content of metabolite is mainly that which is being produced moment-to-moment. The figure is adapted from a plot of data from a computer analysis of a multivariable model.

The kinetics of the generation and elimination of a metabolite relative to those of its drug precursor are important when the metabolite is either toxic or therapeutically active. In the latter instance the kinetics are the kinetics of latentiation. Where the metabolite is toxic, a pattern such as in A would be less likely to generate toxic concentrations as in B.

When the disposition of the drug precursor involves more than one process, or when there is more than one metabolite, the kinetics necessarily are more complex than in the illustrations presented above.

Other Pharmacokinetic Models—Apparent kinetic nonconformities may result when the system does not obey the simple open one- or two-compartment models. In the two-compartment model discussed in this chapter, elimination took place from the central compartment; however,

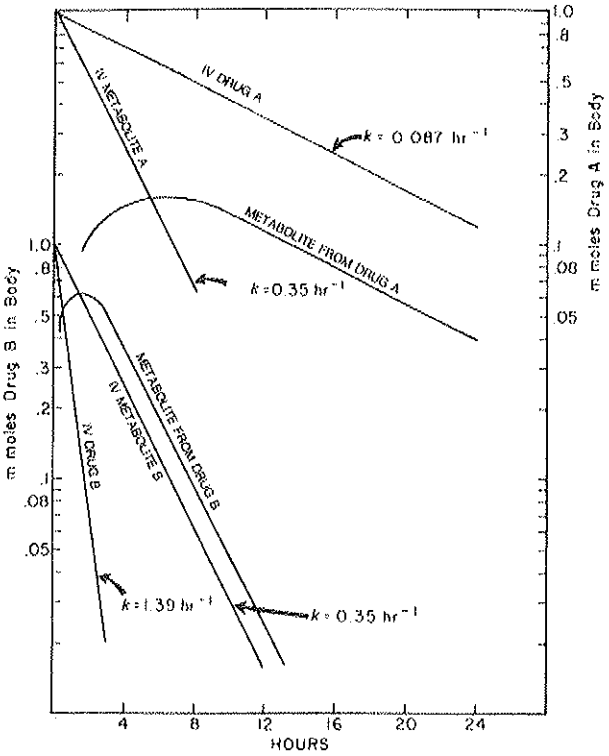


Fig 36-21. Computer plot of the relationship of the amount of drug metabolite in the body to the amount of drug in the body at different relative rates of disposition of drug and metabolite. With Drug A, the metabolite is eliminated at a much faster rate than the parent drug. Curve "IV Metabolite A": the blood concentration when the metabolite is given intravenously; curve "Metabolite from Drug A": the concentration of metabolite actually biotransformed from Drug A. With Drug B the metabolite is eliminated at a much slower rate than the parent drug (courtesy, combined replot of two figures, Martin[10]).

other two-compartment models in which elimination takes place partly or entirely in the peripheral compartment are more appropriate with some drugs. Even absorption into a peripheral compartment appears to occur with some drugs. In addition to alternate two-compartment models, three- or multi-compartment models are required occasionally to account for the pharmacokinetic behavior of certain drugs. In the common *three-compartment* model, the central compartment communicates with two peripheral compartments (which are not interconnected), one called the *shallow compartment* and the other the *deep compartment*. Distribution into the shallow compartment is faster than into the deep compartment.

Many drugs that are described as having one- or two-compartment kinetics actually have more complicated kinetics. There is no drug that displays true one-compartment kinetics, since distribution is never instantaneous. With any drug, sampling within the first minute to one-half hour will show one or more distribution phases.

Nonlinearities—Nonlinearity is a term applied to all nonconformities in which a semilogarithmic plot of plasma concentration-time data cannot be resolved completely into log-linear components, ie, into first-order processes. There may be various causes, such as capacity-limited elimination (ie, saturation of elimination system), capacity-limited absorption or transport, changes in protein binding, changes in pH at the site of absorption, changes in blood flow to the site of absorption and/or elimination, low or erratic dissolution or release rates from dosage forms, low solubility of the drug, drug-induced or other change in body temperature, etc. Some apparent nonlinearities are the result of fitting straight lines to nonlinear data under the assumption that deviations are experimental error.

Protein Binding—The binding of a drug to protein or other macromolecules can affect the pharmacokinetics of a drug, the magnitude of the effect depending on the fraction of the drug that is bound, the fraction of the binding sites that are occupied by the drug and the rates of association and dissociation. If only a small fraction of drug is bound, the kinetic consequences may be minor or negligible, even if binding is very tight. The effect of the binding of a large fraction of drug depends somewhat on whether the drug is bound tightly or loosely; if the rate of dissociation is quite rapid in comparison to the rate of delivery to sites of distribution and elimination or in comparison to the intrinsic rate of elimination, the kinetic consequences also may be minor. The greatest consequences accrue to binding with high capacity and slow dissociation.

It cannot be overemphasized that in the analysis of plasma, the total concentration of drug (ie, both free and bound drug) usually is determined. However, it is only the free drug that can move across cell membranes, and equilibrium or steady-state conditions are established only through the movement of free drug. Therefore, total drug concentrations are defective indicators of a true kinetic situation unless a correction is made for the extent of protein binding. Without such corrections, errors can be serious. Binding to plasma protein has a profound effect not only on V_d but also on apparent renal filtration fraction and clearance, as may be seen in Eqs 13 and 15. If the plasma concentration was not corrected for binding, Cl_{ren} would be in error by a factor of $1/(1 - p)$, where p is the fraction bound; however, when excretion occurs mainly by active tubular transport, protein binding often has a negligible effect on renal clearance. Similarly, when intrinsic hepatic clearance is low, protein binding greatly affects the clearance, the effect being to decrease clearance.

The binding of a drug to plasma proteins retards the rate of distribution and delays the attainment of equilibrium or steady-state conditions. It is as though the transport of some molecules of the drug across a membrane has to wait until these molecules dissociate and are free to diffuse.

When the amount of a drug bound to plasma proteins does not approach saturation, ie, the binding capacity of the proteins, the fraction of drug bound approximately is constant over a therapeutic dose range. However, when the amount exceeds about 50% of the saturation value, the percent of drug bound may vary considerably with dose, which will give rise to dose-dependent kinetics (see below). Under the condition of near-saturation, changes in the protein content of the blood also will make large differences in the percent bound and hence in the various pharmacokinetic parameters. Certain pathological conditions, such as uremia, some congestive heart failure, starvation, etc, may be accompanied by hypoproteinemia and albumin with altered binding properties and hence abnormal pharmacokinetics.

Time-Dependent Kinetics—A drug with low to intermediate intrinsic clearance, and which induces an increase in the activity of its own biotransforming enzyme system, will decrease $t_{1/2}$ and increase clearance and, if its kinetics show two-compartment kinetics, its V_d. Since such an induction requires time, usually several dose-intervals of repetitive dosing, the kinetics vary with time and are called time-dependent. Allosteric (or feedback) inhibition by accumulated metabolites of a drug, or an effect of a drug to impair its route or elimination, also will cause time-dependent (and dose-dependent) changes in the kinetics. Drugs that cause the depletion of some slowly repleteable intermediary factor, such as the depletion of norepinephrine by reserpine or the irreversible inhibition of acetylcholinesterase by isoflur-

ophate, will manifest time-dependent effects on body function which do not correlate with the drug pharmacokinetics. With some drugs, especially central nervous system depressants, the drug effect recruits time-dependent homeostatic counteradjustments that tend to terminate the effect prematurely and to increase the dose requirement for effect (ie, causes tolerance), so that the pharmacokinetics lose their predictability with time. Similarly, drug-induced changes in the receptor properties of the response system will tend to produce a time-dependent dissociation of the pharmacokinetics from the pharmacodynamics.

Dose-Dependent Kinetics—With some drugs, the pharmacokinetics differ more with high, than with low doses. Such changes may be due to: saturation of a biotransforming enzyme or excretory transport system, toxic impairment of the organ of excretion at high doses, differences in intercompartment permeability and V_d at high and low doses, drug-induced changes in blood flow and hence in distribution and clearance, saturation of protein binding sites or the recruitment of new binding sites at high doses, etc. In those instances in which the elimination route is saturated (also called capacity-limited), it is evident that the half-life will increase, as can be seen in Fig 36-22. The cause of the dose-dependent increase in $t_{1/2}$ at the higher doses is the saturation of the enzyme systems that form salicyluric acid and carboxybenzoxyglucuronide. It is usual to speak of the kinetics during the saturation phase as being zero-order, but they are not truly zero-order. The saturated system manifests zero-order kinetics, but alternative routes of elimination, such as through salicyl glucuronide and glomerular filtration and renal tubular secretion, still manifest first-order kinetics, so that elimination is a mixture of zero- and first-order processes. In any event, since elimination is no longer completely a first-order process in the saturation phase, there is no overall elimination rate constant and hence no constant half-life. During repetitive dosing with the large doses, the new $\bar{C}^{ss}$ will be determined by both the zero-order and first-order elimination processes, as well as the dose, but the time required to reach the new plateau will be determined only by the remaining first-order processes; since the first-order overall elimination constant, K, has been diminished, the time-to-plateau will be increased accordingly. Kinetic behavior of this type is mathematically analogous to the familiar Michaelis-Menten expression for enzyme kinetics, and dose-dependent kinetics are sometimes called *Michaelis-Menten* kinetics. They also are called *saturation*, or *capacity-limited*, kinetics.

Examples of important drugs which show dose-dependent kinetics are aspirin, phenylbutazone, probenecid, levodopa, phenytoin and dicumarol. Ethanol obeys essentially zero-order elimination kinetics at blood concentrations above 0.02–0.04%, which is a fact of considerable importance in court cases involving ethanol. The clinical significance of dose-dependent kinetics will be discussed further in Chapter 37.

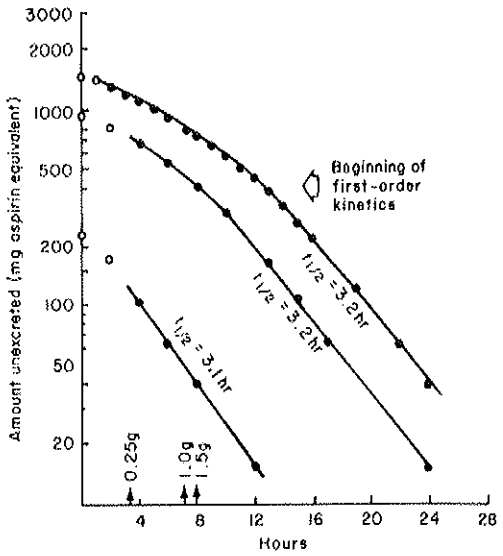


Fig 36-22. Dose-dependent elimination of salicylate in a normal 22-year-old male. Doses taken were 0.25, 1.0 and 1.5 g aspirin, respectively. Vertical arrows on the time axis indicate the time necessary to eliminate 50% of the dose. Stated half-times ($t_{1/2}$) are for straight-line portion of curves where elimination rate is first-order. However, during the early hours after the larger doses, the slope at any time (tangent to the curve) is flatter, hence $t_{1/2}$ is longer, than during the first-order phase (courtesy, Gibaldi and Perrier,[9] modified from Levy[11]).

Chirality—Chiral drugs often are given as racemic mixtures, and the pharmacokinetics and pharmacodynamics of the drugs are studied as if the drug were one entity. It is now becoming clear that this approach may be in error because evidence is accumulating which shows that the pharmacokinetics (as well as pharmacodynamics) of individual enantiomers are not the same and that failure to differentiate among them will give misleading kinetic data for the active form of the molecule. Details of the importance of chirality in pharmacokinetics have been summarized.[12]

Kinetics in the Evaluation of Drugs and Drug Products

The utility of pharmacokinetics in devising appropriate dosage regimens is obvious. Kinetic studies also are important to the study of the influence of inhibitors of elimination, eg, probenecid on the excretion of penicillin, and the effect of one drug on the disposition of another.

Plasma or tissue concentrations and their kinetics are not only valid but essential in comparing the bioavailability of drug products in which the excipients, adjuvants, etc, may vary but the active ingredients are the same. Such data are critical to a proper appraisal of the practice of prescribing drugs by proprietary names.

Kinetics also are employed to compare different drugs, but the meaning of such comparisons is often obscure, and claims of therapeutic superiority based on kinetics must be accepted cautiously. The kinetics of disposition are important to a comparison of drugs in a class in which toxic effects are frequent; it is often desirable to use a drug with a short biological half-life, so that a toxic episode may be terminated quickly upon discontinuation of medication. Furthermore, it is valid to compare the fluctuations in plasma concentration among drugs consequent to multiple-dose administration, provided, of course, that for the class of drugs in question, the extent of fluctuation has an important bearing on efficacy or toxicity.

A comparison of peak or mean blood levels achieved by equal doses of different drugs is not entirely meaningless. It is true that the dose of a drug may be adjusted to compensate for a difference in potency from some reference drug, but it is often difficult for the physician to alter the dose except in multiples of the unit dose provided by the manufacturer. Partly because of the inertia of precedence and habit and partly because it is easier for the physician to memorize

doses as a group, closely related drugs whose potencies differ only moderately may all be available in the same dose. Thus, tetracyclines are available as "250's" or "500's," even though they are not equipotent, sulfonamides as 1 g, etc. It is therefore valid for the physician to choose the drug whose unit dose yields a blood level closest to the optimum. Unfortunately, many physicians do not have the prerequisite knowledge for such a choice and hence may be susceptible to misleading promotional arguments about the superiority of one product over another. Some of these points will be elaborated in the following chapter on *Clinical Pharmacokinetics*.

References

1. Bigger JT: *Am J Med 58:* 479, 1975.
2. Dittert LW: *Drug Intell Clin Pharm 8:* 222, 1974.
3. Truitt EB Jr, *et al: J Pharmacol Exp Ther 100:* 309, 1950.
4. Swintosky JV: *Proc AACP Teacher's Seminar 13:* 140, 1961.
5. Wilkinson GR: *Pharmacol Rev 39:* 1, 1987.
6. Goldstein A, *et al: Principles of Drug Action*, 2nd ed, Wiley, New York, 1974, p 334.
7. Krüger-Thiemer E: In Ariens EJ, ed: *Physico-chemical Aspects of Drug Actions*, vol 7, Pergamon, London, 1968, pp 63–113.
8. Notari RE: *Biopharmaceutics and Pharmacokinetics*, 3rd ed, Dekker, New York, 1980, p 197.
9. Gibaldi M, Perrier D: *Pharmacokinetics*, 2nd ed, Dekker, New York, 1982, pp 30, 276, 441.
10. Martin BK: *Brit J Pharmacol Chemother 31:* 420, 1967.
11. Levy G: *J Pharm Sci 54:* 959, 1965.
12. Walle T, Walle UK: *TIPS 7:* 155, 1986.

Supplementary Reading

Note: Refs 6–9 are textbooks or monographs on general pharmacokinetics. Ref 5 is a monograph on hepatic clearance.

Clark B, Smith DA: *An Introduction to Pharmacokinetics*, 2nd ed, Blackwell, Oxford, 1986.

Curry SH: *Drug Disposition and Pharmacokinetics*, 3rd ed, Blackwell, Oxford, 1980.

Gibaldi M: *Biopharmaceutics and Clinical Pharmacokinetics*, 3rd ed, Lea & Febiger, Philadelphia, 1984.

Gibaldi M, Perrier D: *Pharmacokinetics*, 2nd ed, Dekker, New York, 1982.

Goldstein A, Aronow L, Kalman SM: *Principles of Drug Action*, 2nd ed, Wiley, New York, 301, 1974.

Greenblatt DJ, Shader PI: *Pharmacokinetics in Clinical Practice*, Saunders, Philadelphia, 1985.

Levine RR: *Pharmacology. Drug Actions and Reactions*, 3rd ed, Little, Brown, Boston, 1983.

Niazi S: *Textbook of Biopharmaceutics and Clinical Pharmacokinetics*, Appleton-Century-Crofts, New York, 1979.

Ritschel WA: *Handbook of Basic Pharmacokinetics*, Drug Intelligence, Hamilton, IL, 1976.

Rowland M, Tozer TN: *Clinical Pharmacokinetics: Concepts and Application*, Lea & Febiger, Philadelphia, 1980.

Shargel L, Yu AB: *Applied Biopharmaceutics and Pharmacokinetics*, 2nd ed, Appleton-Century-Crofts, Norwalk CT, 1985.

Wagner JG: *Fundamentals of Clinical Pharmacokinetics*, Drug Intelligence, Hamilton IL, 1971.

Winter ME: *Clinical Pharmacokinetics*, Applied Therapeutics, Spokane, 1980.

CHAPTER 37

Clinical Pharmacokinetics

Douglas E Rollins, MD, PhD
Associate Professor of Medicine and Pharmacology
School of Medicine and College of Pharmacy
University of Utah
Salt Lake City, UT 84112

In Chapter 36 the basic principles of pharmacokinetics were presented. Clinical pharmacokinetics is the discipline in which basic pharmacokinetic principles are applied to the development of rational dosage regimens. In this chapter the concepts of pharmacokinetics are placed into perspective with the development of individualized drug dosage regimens. The clinical significance of the processes of drug absorption, distribution, elimination and influence of disease states on these processes are emphasized. Examples will be given of the ways pharmacokinetic principles can be applied in the calculation and adjustment of dosage regimens designed to fit the pharmacokinetic and pharmacodynamic properties of drugs and specific disease states that alter drug disposition. The principles of therapeutic drug monitoring and the rational use of this clinical science in the management of patients also are discussed.

An individualized dosage regimen for a patient involves a decision about the dose or amount of drug to be administered, interval between doses, route of administration and patient factors that may change during the course of drug administration. The latter implies that there is a plan for monitoring the therapeutic and adverse effects of the drug. Decisions about drug dose, dosage intervals and route of administration are based on the clinical knowledge of the disease being treated, efficacy of the drug in treating the disease and absorption, distribution and elimination of the drug.

Absorption

Drugs are administered by a variety of routes including intravenous, intramuscular, inhalation, oral, rectal, vaginal and topical application to the skin. The choice of the route depends on the many patient- and drug-related factors discussed in Chapter 35. In practical terms, the important considerations in this choice include the systemic availability of a particular dosage form, rate and extent of drug absorption and patient convenience.

Oral Route—This route is chosen most frequently because of ease of administration and patient acceptance. However, the number of variables involved in the absorption of drugs from the stomach and small intestine make the oral route of administration quite complex.

Plasma concentration-time curves will reflect some of these complexities. One of these is the relative rates of absorption of different preparations of the same drug (Fig 37-1[1]), in which preparation *A* represents a simple, rapidly absorbed preparation of a drug; *B* is a more slowly absorbed derivative of the same base. The bioavailabilities of *A* and *B* are identical and *C* is the same compound as *B*, but in a dosage form that is only 50% as bioavailable as *B*. *A* is absorbed rapidly (ie, k_a for *A* is greater than for *B* or *C*) and the peak level is in the therapeutic plasma concentration range.

The advantage of such a preparation is that a pharmacodynamic response can be expected to occur quickly, provided the response is related to plasma concentration. To appreciate the clinical relevance of the situation, consider *A* to be quinidine sulfate, an antiarrhythmic drug. For quinidine sulfate, the absorption rate constant, k_a, is large in relation to the elimination rate constant, k_{el}, and the peak concentration usually occurs in 1 to 2 hr. The rapid absorption is important in clinical situations in which some degree of urgency exists.

It may be desirable, in the initiation of therapy of ominous ventricular premature contractions, to use a preparation with the characteristics of quinidine sulfate. The half-life of quinidine is 4 to 6 hr, so that frequent doses (every 4 hr) are necessary to maintain effective blood concentrations of the drug. The short half-life can be an advantage, since steady-state concentrations of quinidine are achieved within 24 hr (plateau principle). Therefore, one can decide within a day whether quinidine will be useful in suppressing the ventricular premature contractions. However, the fact that a dose must be administered every 4 to 6 hr to maintain therapeutic plasma concentrations is somewhat of a disadvantage in that it is inconvenient and may result in noncompliance.

B, with its slower rate of absorption, reaches a lower peak concentration at a considerably later time even though given in the same dose. There are clinical consequences of this. For example, if *B* was the sustained-release form of quinidine gluconate, it would be less desirable than quinidine sulfate for the initiation of drug therapy, where a rapid therapeutic response is needed. Because of its prolonged

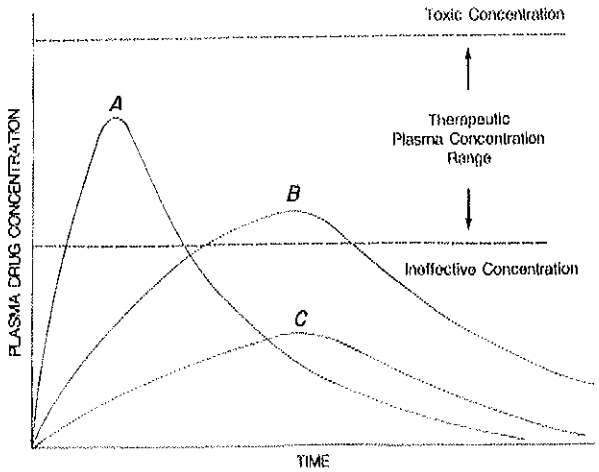


Fig 37-1. Plasma drug concentration-time curves of three preparations of the same drug. *A* is rapidly and completely absorbed. *B* is not absorbed as rapidly as *A* but is 100% available. *C* has the same time-to-peak concentration as *B* but is only 50% as available (Courtesy, adaptation, Benet[1]).

absorption, this preparation commonly is administered every 8 to 12 hr. This is so because the slower rate of absorption enables the dose to be increased commensurate with a longer dose-interval without peak concentrations that rise into the toxic range.

When treating a patient in which a rapid (but not immediate) effect is required (as with asymptomatic ventricular premature contractions), it is advisable to use a dosage form to initiate therapy that is rapidly and completely absorbed. Once the drug is shown to be effective in a particular patient, the dosage form can be changed to one with characteristics similar to *B*, so that less-frequent dosing is required and patient compliance is improved.

The preparation represented by *C in the same dose* as *A* or *B* is probably not an acceptable way to administer this drug. The total amount of drug *C* that is absorbed is only half of that of *B* (area-under-the-plasma concentrations-time curve, AUC, for *C* is half of the AUC for *B*). Thus, it would require twice the dose to attain blood levels equivalent to *A* or *B*.

The treatment of asthma with theophylline is an example in which a rapidly absorbed dosage form is used to initiate therapy and a prolonged-release dosage form is used for maintenance therapy. When a patient has an acute asthma attack or worsening bronchitis that requires bronchodilator therapy, it is advisable to use the theophylline-ethylenediamine complex (aminophylline). This dosage form can be administered either intravenously or orally; the former should be used to initiate treatment in the acute asthmatic patient who requires prompt therapy, so that neither a delay in achieving therapeutic plasma concentrations nor bioavailability are factors in the initial therapeutic response.

Following the administration of a loading dose (see under *Distribution*, page 749), the drug should be given by continuous intravenous infusion until the acute symptoms have subsided, which may take 24 to 72 hr. In the patient with less-severe symptoms, aminophylline can be administered orally four times a day. Once the patient's condition has improved and an effective dose of theophylline has been established, then it may be possible to switch the patient to a prolonged-release formulation for maintenance therapy.

The absorption and bioavailability of Theodur and Sustaire, two sustained-release theophylline preparations, permit 12-hr dosing intervals; Slo-Phyllin Gyrocaps should be given every 8 hr. The total daily dose of theophylline that was required during intravenous aminophylline administration is divided into smaller oral doses given at intervals appropriate for the characteristic of the preparation or dosage form used.

It is important to keep in mind that the absorption and plasma-time curve characteristics for these preparations usually have been established in healthy volunteers or asthmatic patients without other illnesses. Patients who eliminate theophylline rapidly (ie, smokers) may have increased dosage requirements, and the dosage interval may have to be shortened to avoid recurrent asthmatic symptoms between doses.

Prolonged-release dosage forms have the additional advantage that fluctuations in blood levels of the drug will be less than with rapidly absorbed dosage forms. There is evidence for some drugs that the reduction in rapidly changing blood levels may improve efficacy and decrease adverse effects. For example, the dose of fentanyl or ketamine required to maintain anesthesia was reduced by nearly 50% when the drugs were given by continuous infusion rather than by intermittent bolus.[2]

This reduced dose also resulted in more rapid recovery with less-prolonged sedation. These findings suggest that a reduction of fluctuation in the plasma concentrations will reduce total dosage requirement. If such a reduction in plasma concentration fluctuation also applies to oral prolonged-release dosage forms, it would provide a distinct advantage for their use.

The bioavailability of a particular drug product, by any route of administration, can be determined by comparison of the AUC of a drug given by the route of interest with that of the same dose given intravenously (see Chapter 35). In the case of an orally administered drug, it is the ratio of the AUC after an oral dose to the AUC after an intravenous dose. The decreased bioavailability of an oral dose may be due to poor gastrointestinal absorption of the drug because it does not go completely into solution, as it may be degraded in the gastrointestinal lumen, or it does not pass across the intestinal mucosa. Furthermore, in order to reach the general circulation, drugs taken orally must pass through the wall of the gastrointestinal tract and then to the liver via the portal vein. Thus, drug metabolism may occur in the gut wall or in the liver and severely limit the delivery of parent drug to the general circulation.

If the extraction of the drug by the liver is efficient, oral administration results in low bioavailability and sometimes limited pharmacological effect. This is commonly referred to as *first-pass metabolism*. Table I lists some of the drugs known to exhibit first-pass metabolism. Because their extraction is high and their rate of metabolism great, the rate-limiting step in the clearance of drugs in Table I is liver blood flow. The metabolism of these drugs can be referred to as *flow-limited*. The clinical significance of changes in liver blood flow on drug bioavailability will be discussed under *Drug Therapy in Hepatic Disease*.

Different dosage forms of the same drug may have different systemic bioavailabilities. The ratio of the AUC for one dosage form to that of another dosage form is termed the *relative bioavailability*. A drug usually has the highest bioavailability if administered orally as an aqueous solution; finely comminuted drugs in suspension follow closely. However, as a drug is packed into hard gelatin capsules or compacted into tablets, its bioavailability decreases. Furthermore, a drug in one dosage form made by one manufacturer may have a different bioavailability from that of another manufacturer.

With drugs for which bioavailability varies significantly from product to product, if one product initially has been efficacious, it is advisable to continue with that product. If for economical or other reasons the product must be changed to that manufactured by a different company, it is wise to observe the patient carefully for a possible change in clinical response indicative of a change in bioavailability. Products designed for prolonged-release sometimes have a low bioavailability. However, this may not be a problem during maintenance therapy so long as therapeutic serum concentrations are achieved consistently.

The presence of food in the stomach or intestine can have a profound influence on the rate and extent (bioavailability) of drug absorption. Initial absorption studies for a new drug, performed in healthy volunteers, commonly include fasting and nonfasting conditions. Therefore, in general, and for controlled diets, the effect that food may have on

Table I—Drugs that Exhibit First-Pass Metabolism

Acetylsalicylic acid	Metoprolol
Alprenolol	Morphine
Amitriptyline	Nitroglycerin
Desipramine	Nortriptyline
Dopamine	Pentazocine
Imipramine	Prazosin
Isoproterenol	Propoxyphene
Lidocaine	Propranolol
Meperidine	Salicylamide

drug absorption may be known when a drug is introduced into the market. Unfortunately, food-drug interactions are not consistent, and the presence of food may enhance or diminish the absorption of drugs. The most common type of interaction occurs when a food constituent binds the drug and the food-drug complex cannot pass through the gut wall. For example, complexation of tetracycline antibiotics may occur when these drugs are administered with dairy products or with antacids containing aluminum, calcium or magnesium.

The presence of a large meal in the stomach will delay gastric emptying. If a drug that is absorbed in the intestine is ingested with a large meal, the delay in gastric emptying may result in a delay in absorption of the drug. However, the presence of food in the stomach also has been shown to *increase* absorption of some drugs. For example, the bioavailabilities of the β-adrenergic blocking drugs, propranolol and metoprolol, are enhanced by the presence of food.[3] Therefore, because of the difficulty in predicting the absorption pattern of a drug in the presence of food, it is usually advisable to administer drugs when the stomach is empty or 30 min prior to meals; an exception is with drugs which cause gastrointestinal irritation and nausea. These drugs must be given with food to prevent these side effects. It is recommended that such drugs *always* be taken with food to compensate for the differences in absorption that might occur if they were given one time with food and another time without food.

Water taken concomitantly with certain drugs may increase bioavailability. The administration of aspirin, erythromycin stearate, amoxicillin or theophylline with 250 mL of water results in greater bioavailability than if the same drugs are ingested with only 25 mL of water.[4] It is probable that the increased amount of water enhances the amount of drug absorbed by improving drug dissolution as well as by hastening gastric emptying.

Diseases that affect the structure and function of the gastrointestinal tract also are capable of altering the absorption of drugs after oral administration. However, no consistent pattern develops; rather, there appears to be a complex relationship between the effect of the disease on stomach and intestinal functions and the absorption of the drug in question. For example, diseases, such as diabetes mellitus or chronic renal failure, which delay gastric emptying, will markedly delay the absorption and onset of effect of drugs that must reach the small intestine before they are absorbed. This has been a problem with the use of phenytoin in patients with chronic renal failure. Celiac disease and Crohn's disease, which alter the intestinal epithelium, have been studied in detail.[5] In these diseases, absorption of some drugs is affected greatly, but there is no consistent pattern of altered drug absorption.

When a drug is to be administered orally to a patient with altered gastrointestinal motility, diseases of the stomach and small or large intestine, previous stomach or intestinal surgery or gastrointestinal infection, there is a considerable probability that drug-absorption characteristics in these patients will differ from those in healthy volunteers. This may result in a change in the time of peak blood level or the extent of absorption. It is advisable to observe such patients closely for clinical effect during initial drug administration and during chronic dosing in order to assess the influence of alterations in absorption and correct dosing regimens accordingly. Monitoring drug blood concentrations may be beneficial in adjusting dose.

Nonoral Routes—Drugs are administered by a variety of nonoral routes including subcutaneous, intramuscular, intravenous, inhalation, percutaneous, buccal, sublingual, rectal, vaginal, intra-arterial and intrathecal. In the cases of inhalation, topical application to the skin or mucous membranes, rectal, vaginal, intra-arterial or intrathecal administration, the route often is chosen to ensure that drugs reach a specific site with a minimum of systemic absorption. The rationale is that the maximum concentration of drug will be at the site of action so that side effects will be lessened. Nevertheless, if large doses are administered by these routes, enough drug may reach the general circulation to produce side effects. Therefore, the dose and preparation should be such that limited quantities of drug reach the systemic circulation.

The beta-adrenergic agonists, metaproterenol and albuterol, when administered by inhalation, produce bronchodilatation at doses that avoid serious systemic side effects. Similarly, the corticosteroid, beclomethasone, also can be administered by this route for the management of chronic asthma. Low doses of beclomethasone by inhalation are without the serious systemic side effects of oral steroids. However, as the dose is increased beyond two inhalations 4 times a day, for an average daily dose of 400 μg, there is a greater incidence of side effects, including adrenal suppression.

The *topical* administration of drugs rapidly is becoming an important route of drug administration of systemic drugs. Previously used only for the application of drugs for local effects in diseases of the skin, it now is being explored as a means of administering drugs for their systemic effects.

Nitroglycerin commonly is applied to the skin in the form of an ointment or transdermal patches; it is absorbed rapidly and provides sustained blood levels. Sublingual nitroglycerin also is employed to produce therapeutic blood levels; it produces a maximal effect on anginal pain within 3 to 5 min but lasts only 20 to 60 min. In contrast, nitroglycerin ointment provides peak blood concentrations in about 1 hr and the effect on anginal pain may last for several hours. The sublingual tablets should be used to suppress acute angina attacks, whereas nitroglycerin ointment or transdermal patches may be useful to prevent recurrence of episodes of angina for prolonged periods, such as during the night. Whether or not the continuous administration of nitrates by this route will result in the development of tolerance is not clear at this time. Transdermal patches containing clonidine or estrogen are available for the treatment of hypertension or estrogen-replacement therapy, respectively.

Close *intra-arterial* administration of drugs is used to get drugs directly to a target site or organ in high concentration. After it has passed through the target region it is distributed in the entire blood volume, which reduces the systemic levels of the drug and the consequent side effects. One example is the use of cytotoxic drugs for the treatment of primary or metastatic tumors of the liver. The infusion of drugs into the hepatic artery exposes the tumor to higher drug concentrations than can be tolerated with intravenous administration. If the drug is extracted efficiently by the liver, the exposure of sensitive tissues such as bone marrow and gastrointestinal epithelium to the drug will be decreased. For example, after hepatic artery infusion of floxuridine (FUDR), hepatic vein concentrations are 2 to 6 times higher than comparable drug concentrations following intravenous infusion, yet systemic blood concentrations are 75% less.[6] Thus, the therapeutic index of FUDR in the treatment of liver cancer is increased considerably by hepatic arterial infusion. This type of selective drug administration may be beneficial with other drugs that have low therapeutic indices.

Intrathecal injection is used to deliver drugs to the spinal cord or brain in sufficient concentration to produce an effect but at the same time to reduce the incidence or severity of systemic side effects. The intrathecal administration of the cancer chemotherapeutic agent, methotrexate, frequently is employed in the management of leukemic involvement of

the central nervous system. The epidural administration of morphine, which produces long-lasting (6 to 30 hr) analgesia with minimal side effects, is proving to be of benefit in the management of chronic pain.

Distribution

Once a drug is absorbed into the general circulation, it distributes into various tissues and body fluids. The nature and extent of this distribution depends on several factors such as the extent of drug binding to plasma or tissue proteins, blood flow to selected areas of the body, lipid-solubility of the drug and, consequently, its ability to permeate membranes. In clinical practice, concern about drug distribution often arises regarding the penetration of an antibiotic into the central nervous system, into abscesses at any location, into bone for the treatment of osteomyelitis and into specific body fluids such as synovial fluid.

In most cases, the distribution of a drug within the body is determined by the nature of the drug. However, distribution occasionally is altered by the disease process for which it is being used. For example, in healthy individuals, the concentration of penicillin in the nervous system is much less than in serum. However, in patients with inflamed meninges, as in bacterial meningitis, large daily parenteral doses of penicillin can result in bactericidal concentrations in the cerebrospinal fluid. Thus, pneumococcal and meningococcal meningitis can be treated effectively with intravenous penicillin. Increased penetration into the brain in these diseases occurs because the inflamed meninges are more permeable to the penicillin. Also, active transport of penicillin out of the cerebrospinal fluid back into plasma may be impaired in meningitis, thus causing an increase in penicillin concentration in the brain.

In Chapter 36 the term *volume of distribution* (V_d) was introduced. Despite the fact that the V_d of a drug is a very important pharmacokinetic term, it is important to recall that knowing the V_d of a drug does not indicate necessarily how or where a drug is distributed within the body. The abstract nature of the V_d is illustrated with a drug such as the tricyclic antidepressant, amitriptyline. The V_d for amitriptyline is 20 L/kg, which represents a total V_d of 1400 L in a 70-kg man. This large V_d indicates that the amount of drug in the plasma is small in relation to the amount in extravascular compartments and implies that tissue concentrations of the drug probably are very large. Since the volume of total body water in a 70-kg man is less than 70 L, a V_d of 1400 L also illustrates that V_d does not represent a real volume. Drugs with a large V_d usually are distributed extensively to tissues where they commonly are bound to tissue constituents such as DNA or other macromolecules, or dissolved in lipids, whereas drugs that are bound extensively to *plasma* proteins will have smaller V_ds.

One situation in which knowledge of the size of the V_d is useful clinically is in the management of the patient with a severe drug overdose. If a drug such as amitriptyline has a large V_d, it is likely that after an overdose neither hemodialysis nor hemoperfusion will be an effective way of lowering the total body concentration of the drug. Dialysis may lower the plasma drug concentration temporarily, but there will be redistribution from tissues into plasma soon after the dialysis is stopped.

Knowledge of the V_d also is important in determining the loading dose of a drug. This is the dose of a drug administered initially to bring the plasma concentration to a level anticipated during maintenance. An example will illustrate how the V_d is used to determine the loading dose of theophylline. The V_d of theophylline is approximately 0.5 L/kg, and a commonly desired plasma concentration is 10 μg/mL (10 mg/L). Equation 7 (page 728) shows that

$$V_d = \frac{fD}{C_p}$$

where f is the bioavailability factor or the fraction of drug administered that reaches the systemic circulation, D is the dose of drug administered and C is the plasma concentration desired. Since the f for theophylline is 0.96 it can be considered to be 1. Thus

$$0.5 \text{ L/kg} = \frac{1 \cdot D}{10 \text{ mg/L}}$$

and

$$D = 5 \text{ mg/kg} = 350 \text{ mg/70 kg}$$

This dose, administered as a 30-min intravenous infusion, an oral solution or as an uncoated, rapidly dissolving tablet, will result in a peak plasma theophylline concentration of approximately 10 mg/L in patients who have not received theophylline recently.

The V_d usually is considered to be a constant parameter of a drug, so that the loading dose is independent of subsequent changes in drug elimination produced by disease. For example, the loading dose of gentamicin in a patient with severe renal failure usually will not be different from that in a patient with normal renal function. Therefore, therapy can be started with the conventional loading dose without knowing the actual status of renal function.

The severity of renal failure as measured by creatinine clearance (see below) nevertheless will have to be determined prior to calculation of the maintenance dose. There are some clinical situations, however, in which the V_ds of various drugs may be altered so that the loading dose may have to be altered appropriately. The V_d of a drug may be affected by a variety of factors such as protein binding, disease states, body habitus and age. As a rule, the effect of changes in protein binding on the V_d are important only for drugs which are bound 90% or greater to plasma proteins.

Propranolol provides an example in which in patients with chronic liver disease the V_d is increased significantly because plasma protein binding is decreased. This occurs because a greater fraction of unbound drug has access to tissue. The V_d of digoxin in patients with severe congestive heart failure usually is decreased from that in patients with normal cardiac output. Consequently, the loading dose of digoxin is reduced in these patients. Severe dehydration and sepsis result in contraction of the extracellular space and a consequent decrease in the V_d of drugs that largely are confined to this physiological space.

The degree of obesity also may affect the V_d of some drugs. The relative V_d (Δ'; V_d/kg) of water-soluble, lipid-insoluble drugs varies inversely with percent body fat; the Δ' of lipid-soluble, water-insoluble drugs varies directly with body fat. Even in extremely obese patients the increase in body weight may not be accompanied by an increase in the V_d for water-soluble drugs, such as aminoglycoside antibiotics, which will not distribute into fat tissue.

Calculation of the loading dose of these antibiotics in obese patients illustrates this problem. If actual body weight, rather than the ideal body weight or lean body mass, is used to calculate a loading dose of an aminoglycoside antibiotic, elevated peak concentrations may occur in obese patients. Nevertheless, an excessive loading dose is preferable to the risk of possible subtherapeutic concentrations from a miscalculated adjusted dose in a seriously ill patient.

Calculation of maintenance dosing should be made using ideal body weight to avoid consistently elevated peak plasma concentrations. In the first year of life, infants are known to have a larger extracellular space per unit of body weight than adults so that the Δ' of some drugs is also greater. This has been shown to be true for ampicillin, ticarcillin and amika-

cin. Changes in the V_d occur frequently in elderly patients as the result of changes in lean body mass. A linear increase in the Δ' with increasing age has been demonstrated to occur with diazepam.[7]

It should be kept in mind that the V_d for a particular drug in an individual patient may change during therapy. An example might occur when a severely dehydrated patient is treated with intravenous fluids. Unfortunately, there are no accurate means by which the V_d of a particular drug can be determined in an individual patient without first administering the drug in question. Therefore, in situations where one suspects that the V_d may be altered, it is important to monitor blood concentrations of drug, or clinical response, to ensure that therapeutic, and neither toxic nor inadequate, plasma concentrations are being achieved. This particularly is true during initial cumulative drug administration or when a loading dose is being given.

Protein Binding—Pharmacological effect is related closely to the free concentration of drug at its site of action. However, all drugs are bound to some extent to plasma and/or tissue proteins, and the free-drug concentration often may represent only a fraction of the amount of drug in the body. For most drugs the total-drug concentration is measured in plasma and related to an observed therapeutic effect. Thus, recommended therapeutic concentrations commonly are expressed as the total drug concentration in plasma, simply because total-drug concentration is much easier to assay than free-drug concentration. If something occurs that perturbs the protein binding of drug, then either more or less may be free in plasma (and thus free at the site of action) and "standard" therapeutic drug concentration guidelines no longer apply. This situation is made more complex because changes in protein binding may alter elimination as well as distribution. There is definitely a need to understand the therapeutic consequences of alterations in drug-protein binding in order to individualize drug therapy.

The major factors that affect drug-protein binding include the types of proteins available for binding, the binding affinities and capacities and the presence of competing substances, such as endogenous substances and other drugs. Albumin is the major protein in serum, and drug binding to albumin, consequently, has been studied in detail. Drug binding to alpha$_1$-acid glycoprotein and lipoprotein also has been shown to be of clinical significance for certain drugs. There are little data on the ability of other plasma proteins to bind most drugs.

For the purpose of discussing protein binding, drugs can be classified as either acidic or basic (Table II). Acidic drugs commonly bind to plasma albumin, and concomitantly administered acidic drugs may displace one another from their binding sites. Basic drugs may bind to either albumin or alpha$_1$-acid glycoprotein. If a drug is displaced from its binding protein by another drug or by a disease process, the concentration of free drug in plasma (and at the receptor site) will increase temporarily, an effect which then may increase temporarily the pharmacologic response.

The clinical impact of displacement depends on the total amount of drug in the body that is bound, the extent of displacement, whether the drug is also tissue-bound, the V_d and whether the drug is a high-clearance or low-clearance drug. High-clearance drugs are those with an extraction ratio (see below) of close to 1, so that the extraction usually is insensitive to the extent of protein binding. A low-clearance drug, on the other hand, has a lower extraction ratio, and the clearance of the drug may be very sensitive to protein binding.

Warfarin is an example of a low-clearance drug for which the clearance has been shown to vary with the fraction of unbound drug. Thus, after warfarin has been displaced from protein binding sites, $C_{p(free)}$ increases and clearance increases. The increased metabolism will result in the elimination of excess $C_{p(free)}$ and restore the original free-drug levels. Nevertheless, the initial release of bound drug may cause a temporary depletion of clotting factors and consequent bleeding.

The effects of protein displacement are usually of clinical significance only when binding exceeds 85 to 90%. Consider a drug which is 98% bound to plasma proteins. A displacement of 2% potentially will increase free-drug concentration by 100%. However, this does not mean necessarily that free-drug concentration in plasma actually will increase by 100%, because free drug usually distributes quickly into tissues. After redistribution, the actual increase in free-drug concentration in plasma depends on the V_d. If the V_d is large, the increase in plasma concentration may be minimal; if the V_d is small, the concentration at the receptor site may rise significantly and elicit an increase in intensity of drug action. To make matters more complex, a decrease in protein binding also can increase directly the V_d by decreasing the total concentration in plasma, from which the V_d is calculated.

Diseases can alter drug-protein binding by decreasing the amount of protein available for binding and by inhibiting drug binding. Table III lists some conditions that increase or decrease plasma proteins.

Hypoalbuminemia and elevated alpha$_1$-acid glycoprotein have been shown to have the most dramatic effect on drug-protein binding. A normal concentration of serum albumin is 4 g/dL, and a concentration of 2 g/dL would be considered

Table II—Drugs More Than 90% Bound To Plasma Proteins

Basic drugs	Acidic drugs
Alfentanil	Acetylsalicylic acid
Amitriptyline	Cloxacillin
Chlorpromazine	Naproxyn
Desipramine	Penicillin
Diazepam	Phenylbutazone
Flurazepam	Phenytoin
Imipramine	Probenecid
Lidocaine	Sulfinpyrazone
Lorazepam	Tolbutamide
Nifedipine	Warfarin
Nortriptyline	
Propranolol	
Quinidine	
Verapamil	

Table III—Conditions Capable of Altering Plasma Proteins

	Albumin	Alpha$_1$-Acid Glycoprotein
Decreased plasma protein	Burns	Nephrotic syndrome
	Chronic liver disease	
	Cystic fibrosis	
	Protein-losing enteropathy	
	Nephrotic syndrome	
	Pregnancy	
	Chronic renal failure	
	Trauma	
Increased plasma protein	Hypothyroidism	Celiac disease
		Crohn's disease
		Myocardial infarction
		Renal failure
		Rheumatoid arthritis
		Trauma

severe hypoalbuminemia. The effect of hypoalbuminemia on drug-protein binding has the greatest impact if 90% or greater of the drug is bound, if the number of binding sites on albumin are limited or if the drug has a low V_d. It has been shown that a change in plasma albumin concentration from 3.5 down to 2.3 g/dL causes the protein binding of phenytoin to change from 90% to 80.8%.[8] The reduced binding results in an inversely proportional increase in total plasma clearance, so that in steady-state the unbound-drug concentration remains unchanged. Thus, it is probably unnecessary to alter the total daily dose. However, the decrease in total plasma drug concentration poses a potential problem for the interpretation of routine plasma concentrations. This problem is discussed in further detail under *Drug Therapy in Renal Disease*.

Diseases also can affect the affinity of drugs for albumin. The best-known example occurs in chronic renal failure, in which accumulated endogenous compounds, which are not significantly removed by dialysis, displace acidic drugs from albumin binding sites. In disorders or situations in which free fatty acid levels are increased, acidic drugs are displaced from albumin binding sites. Quantitatively, when the free fatty acid/albumin ratio exceeds 3.5, the binding of acidic drugs usually is reduced significantly.[9]

Elimination

The elimination of drugs from the body usually occurs either by excretion into the urine or by biotransformation to metabolites that are eliminated in the urine or feces. The mechanisms whereby the kidneys and liver eliminate drugs and the pharmacokinetic principles behind these processes were presented in Chapters 35 and 36, respectively. In this section, emphasis will be placed on the practical application of these principles toward the development of individualized dosage regimens.

When drugs are approved by the FDA, their elimination has been studied in detail, usually only in healthy volunteers. Nevertheless, there is often enough information available to make rational decisions about the individualization of drug doses in patients who might have impaired elimination. The most important information is whether the drug is eliminated unchanged in the urine or biotransformed in the liver. With a drug for which the major route of elimination is renal, it is necessary to know if excretion is by tubular secretion, glomerular filtration or by a combination of secretion and filtration. With a drug of which the elimination is principally by the liver it is necessary to know if the biotransformation is primarily by a Phase I (oxidation) reaction or a Phase II (conjugation) reaction, if the metabolite(s) is/are pharmacologically active and if the drug exhibits first-pass metabolism. With the knowledge of these facts about each drug, one can determine if it is necessary to adjust the dosage regimen in a patient with kidney or liver impairment.

As indicated in Chapter 36, drug clearance is a more direct expression of elimination than is half-life. This is mentioned here only to remind the reader to be cautious about equating impaired renal or hepatic function with a change in drug half-life. If a decrease in the renal elimination of a drug is accompanied by an increase in half-life, it is necessary to know this to adjust the dosage regimen. However, the elimination half-life of a drug is a complex function of elimination and the V_d, and it is possible to have a change in the V_d in patients with renal or hepatic impairment such that there is no alteration in half-life. Furthermore, it is possible to have a drug with a high total body clearance yet a long half-life. This seeming contradiction occurs when drugs with a very high clearance also have a large V_d.

One class of drugs that displays this contradiction is the tricyclic antidepressants; the members have rapid clearances of about 1500 mL/min as the result of hepatic metabolism, but their plasma elimination half-life may be as long as 20 hr. Because of their large V_d (1000 to 2000 L) and rapid redistribution between tissues and plasma, drug cleared from the plasma almost completely is replaced by drug from the peripheral compartment. As already mentioned, this is important to remember when deciding about the use of extracorporeal (hemodialysis or hemoperfusion) systems to remove drugs from the body of an overdosed patient.

For a drug with a half-life of 20 hr it might appear that an extracorporeal system would enhance drug elimination. However, clearance of the tricyclic antidepressants by dialysis is small compared to normal hepatic clearance. If the drug also has a large V_d, redistribution likely would keep the plasma levels elevated and hemodialysis or hemoperfusion would have to be continued for an unusually long time to enhance significantly the removal of drug from the body.

Renal Excretion—Unchanged drug or drug metabolites can be eliminated from the body by way of the kidneys, as mentioned above. Drug excretion by this route takes place either as a result of filtration through the glomerulus, by tubular secretion or both. A knowledge of how a drug is excreted can be useful in predicting the effect that renal disease will have on its elimination. Drugs that are excreted by tubular secretion generally can be divided into organic acids, such as penicillin and probenecid, and organic bases such as cimetidine.

As indicated in Chapter 35, the organic acids and bases are secreted by separate transport systems. Among the organic acids there is competition in transport such that the coadministration of two such drugs can result in decreased elimination and elevated blood concentrations of each.

Sometimes this competition can be used to advantage, as in the administration of probenecid in combination with penicillin in the treatment of gonorrhea. The result is that the clearance of penicillin is reduced and the plasma penicillin concentrations remain high for a prolonged period of time; the combination is more effective than penicillin alone. Since the therapeutic index of penicillin is high, such interactions are useful. However, if probenecid is administered with the cytotoxic drug, methotrexate, the secretion of the latter drug is impaired and significant toxicity may occur. When tubular secretion is high, plasma protein binding usually does not affect active secretion by the proximal tubule.

Most drugs are excreted by the kidney via filtration across the glomerular membrane. Glomerular filtration is a passive, nonsaturable process. Because of the small size of the pores of the glomerular membrane, only free drug in plasma can be filtered; consequently, drugs that are bound to plasma proteins are filtered poorly. Displacement from proteins actually can increase the amount of drug filtered in the glomerulus and hence eliminated in the urine.

The glomerular clearance of drugs is directly proportional to the glomerular filtration rate (GFR). It follows that a decrease in GFR will result in a proportional decrease in the rate of glomerular elimination of a drug. Thus, measurement of the GFR can be very helpful in the individualization of dosage regimens in patients with impaired renal function. The GFR generally is estimated by measuring the clearance of either inulin or creatinine. Inulin must be infused intravenously, whereas creatinine, a product of muscle metabolism, is released *in vivo* at a relatively constant rate, thus obviating the need for constant intravenous infusion. Urinary creatinine excretion usually exceeds the amount filtered by about 10% because of a small amount of renal tubular secretion of creatinine. However, because determination of GFR by creatinine clearance is inexpensive and easy to do and, because the difference between inulin and

creatinine clearance is not significant *clinically*, creatinine clearance commonly is used to estimate GFR. It is very important to realize that the creatinine clearance is an accurate estimate of GFR only if renal function is stable. If renal function is decreasing, serum creatinine concentrations will be increasing, and it may take several days to reach a new steady-state. Until a new steady state is reached, the GFR cannot be estimated accurately from serum creatinine concentrations, and serum creatinine should not be used to calculate an individualized dose of a drug. Although creatinine clearance only measures the GFR, it frequently is used in the determination of the dosage regimens of drugs that are eliminated both by filtration and by tubular secretion. Unfortunately, there is no simple test to measure tubular secretion. Therefore, dosage adjustment based on creatinine clearance may not be appropriate for patients receiving drugs that are secreted actively by the renal tubules.[10]

The effect of changes in urine pH and urine flow on drug excretion already have been discussed in Chapter 35. In routine drug therapy, these parameters are not considered to be of great importance. However, the alkalinization of urine to pH 8 by the administration of sodium bicarbonate is used routinely to treat overdoses of phenobarbital and salicylates, since ionization of these weak acids reduces their reabsorption and increases their elimination.

Drug Therapy In Renal Disease—Drug administration to patients with impaired renal function is complicated by their associated medical problems, by the number of drugs they receive and by the alterations in drug disposition and elimination that occur. In renal disease, the protein binding of acidic or neutral, but not basic, drugs in plasma usually is altered. Some of the reasons to explain changes in protein binding include:

1. Hypoalbuminemia that occurs as a result of protein loss in the urine.
2. Competition for protein binding sites with small acidic molecules that accumulate in uremia.
3. Changes in the conformation of albumin that results in decreased affinity for binding sites.
4. Accumulation of drug metabolites that might displace parent drug from proteins.

Whichever the cause for changes in binding, the clinical importance of changes in plasma binding and/or protein concentration is that care must be used to interpret plasma drug concentrations.

Measured plasma drug concentrations usually are reported as total drug, ie, bound plus free drug. For example, therapeutic plasma concentrations of phenytoin in persons with normal plasma protein content are 10 to 20 mg/L, of which only 1 to 2 mg/L represents free drug. In patients with renal failure, the *free* phenytoin concentration is unchanged, whereas the *total* drug concentration falls to 5 to 10 mg/L, because of changes in protein concentration. The clinician might, therefore, be mislead into thinking that an increase in dose was necessary to increase the plasma concentration. In fact, because the free phenytoin levels are unchanged in patients with renal disease a dosage adjustment is not warranted. The renal elimination of metabolites can also be affected by impaired renal function.

The uremic state has been shown to have an effect on the biotransformation of many drugs. However, the effects of uremia on drug metabolism often are inconsistent and not predictable, and the clinical significance of such effects usually are not known. The clinical importance of the reduced elimination of drug metabolites is better understood. Table III in Chapter 36 lists active drug *metabolites*, many which are eliminated by the kidneys.

Procainamide is acetylated in the liver to *N*-acetylprocainamide, which has cardiac effects similar to those of the parent drug. This metabolite is eliminated by the kidneys, and its plasma concentration is increased in patients with impaired renal function. Patients with renal failure who are treated with procainamide should be observed closely for signs of clinical procainamide toxicity, and plasma concentrations of both procainamide and *N*-acetylprocainamide should be monitored.

Dosage adjustment of drugs in patients with renal impairment should be based on a knowledge of the pharmacokinetic parameters of the drug and, when indicated, on monitoring of plasma drug concentration. The aim of individualizing dosing regimens in patients with impaired elimination (renal or hepatic) is to maintain an average plasma concentration ($C_{p(ave)}$) similar to that of patients with normal elimination and, thus, to avoid unnecessary toxicity or loss of efficacy.

In Eq 32 in Chapter 36 it can be seen that $C_{p(ave)}$ is a direct function of dose (D) and bioavailability (f) and an inverse function of the dosing interval (τ) and clearance ($V_d \cdot k_el$). In the patient with impaired elimination or decreased clearance, $C_{p(ave)}$ will increase until a new plateau is reached (plateau principle). If clearance is impaired markedly or if the therapeutic index of the drug is small, toxicity may occur.

It is apparent from the same equation that either an appropriate decrease in dose or increase in the dosing interval will offset a decrease in elimination, and a $C_{p(ave)}$ can be attained that is similar to that in a nonimpaired patient.

In the patient with renal impairment, individualization of drug therapy requires knowledge of the degree of impairment and its effect on drug elimination in order to choose a proper dose or dosing interval to achieve a desired $C_{p(ave)}$. As discussed above, the endogenous creatinine clearance is usually the most practical index of GFR and it is used widely (with the limitations indicated) to determine the degree of renal impairment in a patient with renal disease.

The translation of the degree of impairment into a dosage regimen is not simple. In the literature there are a variety of nomograms and equations available to aid in calculating dosage regimens in patients with renal impairment. Each has its proponents and opponents and each is based on a set of assumptions that provide limitations to its use. None take into account all of the complexities discussed above. Therefore, a nomogram or an equation used to determine a dose of a drug to be given to a patient with renal impairment must be used only as a guideline and, when possible, should be used along with monitoring of plasma drug concentration, when indicated, and careful clinical observation to ensure optimal therapy.

Drug clearance in patients with renal insufficiency (Cl_{ri}) can be estimated from the relationship of the creatinine clearance in the renal-impaired patient, the creatinine clearance of normal persons and the clearance of drug by renal and nonrenal clearance mechanisms according to the equation

$$Cl_{ri} = Cl_{renal} \times \frac{Cl_{creat\ impaired}}{Cl_{creat\ normal}} + Cl_{nonrenal} \qquad (1)$$

where Cl_{renal} is the normal renal clearance, $Cl_{creat\ impaired}$ is the creatinine clearance in the patient, $Cl_{creat\ normal}$ is the creatinine clearance in normal persons and $Cl_{nonrenal}$ is the nonrenal clearance. The renal and nonrenal clearances may not be available; therefore, to determine a proper dosage regimen, one must rely on the pharmacokinetic information that is available in the literature; the elimination rate constants, k_el, in normal patients and in patients with complete anuria frequently are available. The values for these constants for many drugs have been listed in Table IV. Dettli[11] has derived a nomogram in which these elimination rate constants and the creatinine clearance can be used to determine an individualized dosage regimen for patients with

Table IV—Drug Elimination Rate Constants in Normal and Anephric Patients

Drug	Normal k_{el} (hr^{-1})	Anephric k_{el} (hr^{-1})
Alpha-methyldopa	0.17	0.03
Amikacin	0.40	0.04
Amoxicillin	0.70	0.10
Amphotericin B	0.04	0.02
Ampicillin	0.70	0.10
Carbenicillin	0.60	0.05
Cefazolin	0.40	0.04
Cephacetrile	0.70	0.03
Cephalexin	1.00	0.03
Cephalothin	1.40	0.04
Cephaloridine	0.50	0.03
Chloramphenicol	0.30	0.20
Chlorpropamide	0.02	0.008
Chlortetracycline	0.10	0.10
Clindamycin	0.47	0.10
Cloxacillin	1.40	0.35
Colistimethate	0.20	0.04
Digitoxin	0.004	0.003
Digoxin	0.017	0.006
Doxycycline	0.03	0.03
Erythromycin	0.50	0.14
Ethambutol	0.58	0.09
Fluorocytosine	0.24	0.01
Gentamicin	0.30	0.01
Isoniazid		
(fast acetylators)	0.60	0.20
(slow acetylators)	0.20	0.08
Kanamycin	0.40	0.01
Lidocaine	0.40	0.36
Lincomycin	0.15	0.06
Methicillin	1.40	0.17
Minocycline	0.05	0.03
Nafcillin	1.20	0.48
Oxacillin	1.40	0.35
Oxytetracycline	0.08	0.02
Penicillin G	1.40	0.05
Polymyxin B	0.16	0.02
Procainamide	0.22	0.01
Propranolol	0.20	0.16
Quinidine	0.07	0.06
Rifampin	0.25	0.25
Streptomycin	0.27	0.01
Sulfadiazine	0.08	0.03
Sulfamethoxazole	0.70	0.70
Tetracycline	0.08	0.01
Ticarcillin	0.60	0.06
Tobramycin	0.36	0.01
Trimethoprim	0.60	0.02
Vancomycin	0.12	0.003

decreased renal function. This nomogram is reproduced in Fig 37-2.

An example of how this nomogram can be applied is as follows. The ratio $k_{el(anephric)}/k_{el(normal)}$ is the fraction of the usual dose of a drug to be administered when there is anuria. When this ratio is entered on the left ordinate of the nomogram in Fig 37-2 and connected by a line to the upper-right-hand corner, the dose fraction is described for a range of creatinine clearances from 0 to 100 mL/min (100 mL/min is that of a normal 70-kg person). A line then is drawn vertically from the patient's creatinine clearance on the abscissa to the dose fraction line. From this point of intersection, a second line is drawn horizontally to the left ordinate of the nomogram. The point of intersection on the left ordinate is the dose fraction for that particular drug corresponding to the compromised creatinine clearance.

Insofar as the maintenance dose is concerned, the dosage regimen in the patient in renal failure can be modified by

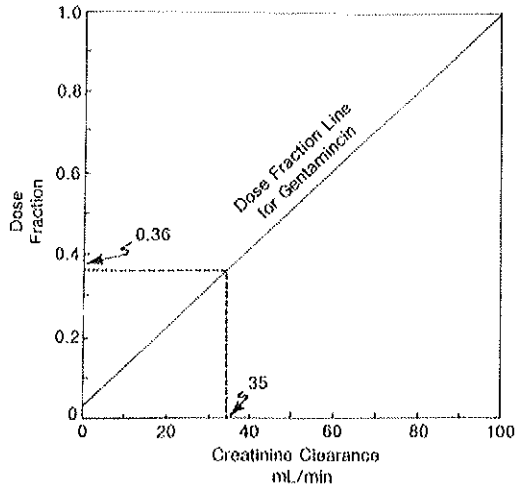


Fig 37-2. Nomogram used to determine the fraction of a dose that should be administered to a patient with a particular creatinine clearance. An example is given for a patient with a creatinine clearance of 35 mL/min and a ratio of $k_{el(anephric)}/k_{el(normal)}$ of 0.03. The dose fraction in this case is determined to be 0.36. This dose fraction then is used to adjust the dose or dosage interval for a patient with that degree of renal impairment (courtesy, adaptation, Dettli[11]).

adjusting either the dose or the dosage interval according to the calculated dose fraction. The maintenance dose can be adjusted by multiplying the normal dose by the dose fraction

$$D_{ri} = D\text{-Dose Fraction} \quad (2)$$

where D_{ri} is the dose in renal insufficiency, D is the usual dose in normal persons and dose fraction is the value determined from the nomogram as described above. The dosage interval, τ, can be adjusted by dividing by the dose fraction

$$\tau_{ri} = \tau/\text{Dose Fraction} \quad (3)$$

where τ_{ri} is the dosage interval in renal insufficiency. An example of an adjustment in a gentamicin dosage regimen for a patient with an impaired creatinine clearance of 35 mL/min is as follows: the usual gentamicin dosage regimen in a patient with normal renal function is a loading dose of 80 mg followed by 80 mg every 8 hr. From Table IV it can be seen that

$$k_{el(anephric)}/k_{el(normal)} = 0.01/0.30 = 0.03$$

When 0.03 is entered on the left ordinate of the nomogram and a line is extended to the upper-right-hand corner, the dose-fraction line for gentamicin is described. From a creatinine clearance of 35 mL/min on the abscissa a line is drawn vertically to the gentamicin dose-fraction line. From this point of intersection a corresponding point on the left ordinate of the nomogram is a dose fraction of 0.36. The dosage interval then can be adjusted as

$$\tau_{ri} = \tau/\text{Dose Fraction}$$
$$= 8 \text{ hr}/0.36$$
$$= 22.2 \text{ hr}$$

Thus, in a patient with such an impaired renal function, a once-a-day dose of 80 mg is likely to maintain therapeutic plasma concentrations. The maintenance dose for gentamicin in this patient also could be adjusted using Eq 2 as follows

$$D_{ri} = D\cdot\text{Dose Fraction}$$
$$= 80\text{ mg}\cdot 0.36$$
$$= 28.8\text{ mg}$$

Thus, 29 mg administered every 8 hr would provide therapeutic plasma concentrations in this patient. The decision to adjust the dose or the dosage interval also should be individualized. Fluctuations in plasma concentration of gentamicin will be less if the dosage interval is lengthened to 24 hr. However, there may be a therapeutic reason to have peak plasma concentrations occur 3 times a day rather than only once. As mentioned above this, or any other nomogram or calculation for dosage adjustment, is only an approximation. Once the dosage adjustment has been made, careful clinical observation and, when indicated, monitoring of plasma concentrations is warranted. Since the loading dose depends primarily on the V_d, a change only in k_{el} does not necessitate a change in the loading dose.

Drug Therapy in Hepatic Disease—The biotransformation of drugs is discussed extensively in Chapter 35. Although many organs are involved in drug biotransformation, the liver is the most important. One might therefore assume that all patients with liver disease would demonstrate a predictable decline in drug elimination by the liver. This is not the case. There are several factors that complicate the management of drug therapy in patients with liver disease.

There are no routinely performed laboratory tests that predict the effect of liver disease on drug metabolism. Unlike the correlation between creatinine clearance and renal clearance of drugs, there is not a good correlation between the commonly available tests of liver function and drug clearance by the liver. In fact, the elimination rates of many biotransformed drugs are unaffected by liver disease.

Drug elimination by the liver may be affected by several factors including liver blood flow, protein binding and volume of distribution, in addition to drug-metabolizing capacity.

Liver disease is not a single well-defined entity but comprises a number of various structural and functional alterations. These include inflammation and necrosis, which generally alter only liver cell function and hence drug-metabolizing activity; cirrhosis, which may impair both liver cell function and liver blood flow; cholestasis, which may impair both biotransformation and biliary elimination and neoplasia, which may both impair cell function and decrease blood flow.

The discussion of biotransformation in Chapter 35 indicates that the process of hepatic elimination of drugs is complex, involving many different types of chemical reactions. While this is true, for practical purposes it is most important to know whether a drug is metabolized by an oxidation (Phase I) or conjugation (Phase II) reaction. The specific type of chemical reaction is of less clinical importance. Many drugs are biotransformed first by an oxidation reaction and the resulting metabolite then is conjugated to facilitate urinary excretion. In these cases it is the oxidation reaction that probably is most important.

The clinical significance of knowing the general reactions involved in the metabolism of drugs is related to administration of such drugs in the patient with hepatic impairment. It generally is accepted that liver disorders which affect hepatocyte cell function will impair drug oxidation long before drug conjugation is altered. A specific example occurs within the benzodiazepine class of drugs. On the one hand, chlordiazepoxide and diazepam are metabolized initially by oxidation reactions that have been demonstrated to be impaired in patients with alcoholic cirrhosis.[7,12]

Accordingly, the elimination of these drugs is decreased, and elevated blood levels may result during chronic therapy. On the other hand, oxazepam and lorazepam undergo only conjugation with glucuronic acid prior to being eliminated in the urine. Glucuronidation does not appear to be affected in clinically stable alcoholic cirrhosis, and the elimination of these drugs is no different than in healthy volunteers.[13,14] From a pharmacokinetic point of view, oxazepam and lorazepam are more rational choices than diazepam or chlordiazepoxide for use in patients with alcoholic cirrhosis.

Most studies of drug elimination in patients with liver disease have been performed in patients with either acute viral hepatitis or alcoholic liver disease. One should be careful about extrapolating these data to patients with other types of liver disease, such as chronic forms of hepatitis, neoplasias of the liver or cholestasis. Furthermore, one must not extrapolate studies of the metabolism of one drug in patients with liver disease to another drug, even though the metabolic reactions appear to be similar. There is a multiplicity of subpopulations of cytochrome P-450 enzymes. One drug may be metabolized by one of these subpopulations, while another drug is metabolized by another enzyme. For this reason, there is often poor correlation between the oxidations of two drugs.

Hepatic disease also can produce changes in serum proteins and in liver blood flow which can influence the elimination of drugs. Because the liver is the site of synthesis of serum proteins, patients with severe chronic liver disease frequently have decreased protein binding of drugs. In addition, there may be decreased protein binding as a result of qualitative changes in serum proteins. Liver blood flow is dominated by the portal venous system that drains the mesenteric veins. Thus, all drugs absorbed from the oral route pass through the liver via the portal vein. In certain types of liver disease, most commonly alcoholic cirrhosis, there is shunting of the portal circulation away from functioning hepatocytes. This leads to increased pressures within the portal system and shunting of drugs away from the drug-metabolizing enzymes.

One method of classifying drugs by the characteristics of hepatic elimination is to divide them into those with a high hepatic extraction ratio and those with a low hepatic extraction ratio. As described in the explanation of Eq 23 of Chapter 36, the hepatic extraction ratio is defined as

$$E = \frac{C_{ap} - C_v}{C_{ap}}$$

where C_{ap} is the hypothetical mean of mixed hepatic arterial and portal venous drug concentrations, and C_v is the hepatic venous drug concentration. The hepatic clearance, Cl_H, of a drug is determined by its extraction ratio as

$$Cl_H = HBF\cdot E$$

where HBF is total hepatic blood flow. The classification of drugs according to their hepatic extraction ratios is shown in Table V. Hepatic blood flow is usually the rate-limiting factor in the hepatic clearance of drugs with high extraction

Table V—Classification of Drugs According to Their Hepatic Extraction Ratios

Drugs with an Extraction Radio Greater than 0.5	
Lidocaine	Nortriptyline
Propranolol	Morphine
Pethedine	Labetalol
Pentazocine	Verapamil
Propoxyphene	Metoprolol

Drugs with an Extraction Ratio Less than 0.5	
Binding-Sensitive	*Binding-Insensitive*
Phenytoin	Theophylline
Diazepam	Acetaminophen
Tolbutamide	Hexobarbital
Warfarin	Chloramphenicol
Chlorpromazine	
Digitoxin	
Quinidine	

ratios, and the metabolism of such drugs are considered to be *flow-limited* metabolism. These drugs demonstrate first-pass metabolism in that after oral administration a major portion of the drug does not reach the systemic circulation. Their bioavailability is low and their metabolism is sensitive to anything that alters hepatic blood flow. Thus, for example, the elimination of lidocaine can be decreased substantially in patients with congestive heart failure, which usually causes a reduction in hepatic blood flow. In patients with cirrhosis and portal hypertension, the shunting of blood away from functioning hepatocytes has the greatest impact on drugs with a high hepatic extraction ratio. In patients with portal hypertension, the bioavailability of drugs with a high extraction ratio may be increased significantly, so that toxic blood levels may result. At the present time there is no routine laboratory test that will predict this effect in an individual patient. Rather, it is advisable to start with a low dose of drug and increase the dose slowly to achieve the desired response.

The rate of metabolism for drugs with a low extraction ratio is dependent on the concentration of drug at the hepatic enzyme site, which is proportional to the free concentration of drug in plasma. Consequently, drugs in this class can be divided further into those in which hepatic elimination is either sensitive or insensitive to protein binding. Drugs with a hepatic elimination distinctly sensitive to protein binding are generally 80 to 99+% bound, whereas drugs with a hepatic elimination clearly insensitive to protein binding are less than 30% bound. Conditions that affect plasma protein binding can have a significant effect on the hepatic clearance of a binding-sensitive drug but usually not a binding-insensitive drug.

Although much is known about the hepatic metabolism of drugs and the factors that can affect their hepatic elimination, the use of drugs in patients with potential altered hepatic clearance is still empirical in that there are no specific guidelines relating the severity of hepatic disease and drug elimination. To a great extent this is due to the multiplicity of drug-metabolizing enzymes, and it is unlikely that a single or simple battery of laboratory tests will suffice to predict the hepatic elimination of all drugs. Applying the known facts about liver disease along with the knowledge of drug elimination by the liver usually will permit a rational use of drugs in patients with disorders of the liver.

Therapeutic Drug Monitoring

Rational drug therapy requires individualization of the dosage regimen for a particular patient. In many instances this can be done by *monitoring the clinical response* to drug therapy. For example, if a patient with hypertension is not responding to therapy and there is no reason to suspect poor compliance, it may be appropriate to increase the dose until the patient's blood pressure is under control. Whenever a drug is administered, well-defined therapeutic end-points should be a preferred part of the management plan.

Observation of the clinical response or monitoring a reliable laboratory test may be easy with certain classes of drugs such as antihypertensives, oral hypoglycemics, oral anticoagulants, analgesics or drugs used to lower serum uric acid or serum lipids. For other drugs, the definition of a therapeutic end-point may not be clear or the onset of toxicity may occur at dosages only slightly above therapeutic concentrations. For some of these drugs one should monitor the serum drug concentration and thus determine if the dose administered to an individual patient is achieving therapeutic concentrations.

The following are several criteria and typical examples that should be considered before measured drug serum concentrations are of clinical value.

The drug must have a reversible action. An example of drugs with irreversible action would be the alkylating agents which exert a lasting effect after a single dose. At the present time there seems to be little need for routinely monitoring the plasma concentration of these drugs.

The development of tolerance at the receptor site should not occur. A therapeutic concentration range for morphine is not rational, since the dose requirements may increase with use.

The pharmacokinetic properties of the drug are taken into account in the blood sampling schedule. If sampling is performed in a maintenance regimen, steady state should have been achieved prior to sampling. Steady state may occur 4 to 5 half-lives after the initiation of therapy if a loading dose is not administered. Changes in drug half-life produced by disease must be taken into account. Qualitative differences in the metabolism or excretion of drugs also are known to occur in patients with hepatic and/or renal disease. For example, patients with impaired renal function may experience prolonged respiratory depression when treated with morphine, due, in part, to the accumulation of an active metabolite, morphine-6-glucuronide. For drugs with a short half-life, peak (1 or 2 hr after oral dosing) and trough (predosing) determinations are advisable. The distribution phase should be complete before drug concentrations are measured. Slow-release formulations of drugs have different absorption characteristics and different plasma concentration versus time profiles that must be taken into account when interpreting a single plasma concentration. The chronic administration of some drugs (ie barbiturates) results in the induction of hepatic drug-metabolizing enzymes. A decrease in the steady-state plasma concentration of that drug, or others metabolized by the induced hepatic enzymes, may occur unless the dose of that drug is increased.

The presence of active metabolites should be taken into consideration. The serum concentrations of the *N*-acetylprocainamide metabolite of procainamide should be considered when assessing antiarrhythmic activity after administering procainamide. This is particularly true in patients with renal failure who may eliminate the metabolite slowly. Active metabolites also are responsible for toxicity (ie acetaminophen). Most assays for the measurement of plasma drug concentrations do not account for active toxic metabolites that are present at very low plasma concentrations.

The analytical method must be sensitive enough to measure accurately the expected serum concentrations and selective enough to be certain that interfering substances will not influence the results. Most clinical drug assays do not distinguish between enantiomers if a racemic mixture of drug is administered. It is important to consider this when interpreting the plasma concentration of a drug if one enantiomer is more active or there is stereoselective disposition. The (*S*)-warfarin enantiomer is about five times more potent in man than the (*R*)-enantiomer; the *S*-(+)-enantiomer of disopyramide is bound more avidly to plasma proteins than its corresponding *R*-(−)-enantiomer. Some drugs (ie phenytoin) may be adsorbed by plastics in intravenous tubing, syringes and blood-collection tubes. When analytical results do not fit the clinical situation, consideration should be given to adsorption as a potential problem.

The data must be evaluated in the context of sound clinical judgment. Treat the patient, not the serum drug concentration. An example is the patient who is taking digoxin and develops a low plasma potassium. Hypokalemia makes the myocardium more sensitive to the rhythm disorders produced by digoxin. Thus, the patient with a normal serum digoxin concentration may experience drug-induced cardiotoxicity if hypokalemia also is present.

Table VI—Therapeutic Ranges for Drugs

Amikacin	Trough	4–8	mg/L
	Peak	20–30	mg/L
Carbamazepine		4–8	mg/L
Digoxin		0.8–2	μg/L
Disopyramide		2–5	mg/L
Ethosuximide		40–100	mg/L
Gentamicin	Trough	0.5–2	mg/L
	Peak	5–10	mg/L
Lidocaine		1.2–5	mg/L
Phenobarbital		15–40	mg/L
Phenytoin		10–20	mg/L
Primidone (see phenobarbital)		5–12	mg/L
Procainamide		4–10	mg/L
N-Acetylprocainamide		10–30	mg/L[a]
Quinidine		1.5–4.5	mg/L
Theophylline		10–20	mg/L
Tobramycin	Trough	0.5–2	mg/L
	Peak	4–10	mg/L
Valproic Acid		50–100	mg/L

[a] Total of procainamide and *N*-acetylprocainamide.

Table VII—Pharmacokinetic Parameters of Commonly Monitored Drugs

Drug	Volume of distribution (L/kg)	Protein binding (%)	Oral availability (%)	Route of elimination	Half-Life		Dose adjustment required	
					Normal	Anephric	Renal failure	Liver failure
Amikacin	0.25	<5	Parenteral only	Renal	3 hr	2–4 days	Yes	No
Carbamazepine	0.8–1.4	75	70	Hepatic—epoxide metabolite is active	10–26 hr	—	No	No
Digoxin	5.1–7.4	20–40	50–93	Renal	33–51 hr	3.6 days	Yes	No
Disopyramide	0.5	50–80	80–85	Renal and Hepatic	6–10	45	Yes	No
Ethosuximide	0.62	Negligible	100	Hepatic	60 hr adults 30 hr children	—	No	No
Gentamicin	0.25	<5	Parenteral only	Renal	2 hr	2–3 days	Yes	No
Lidocaine	1.6	60	Parenteral only	Hepatic—metabolites are active	1.5 hr	—	No	Yes
Phenobarbital	1.0	46	80–100	Hepatic primarily	3–4 days	—	No	Yes
Phenytoin	0.6	90	90	Hepatic	10–30 hr concentration dependent	—	No	Only in severe cases
Primidone	0.6	14	100	Hepatic—phenobarbital and phenylethylmalonylamide (PEMA) are active metabolites	3–12 hr 29–36 hr metabolites	—	No	No
Procainamide	2.2	15	75–95	Renal and Hepatic *N*-acetylprocainamide is active	2.5–4.5 hr	10–15 hr	Yes	No
Quinidine	0.5	60–80	70–95	Hepatic—metabolite active	6 hr	—	No	No
Theophylline	0.3–0.6	55	Complete	Hepatic	3–9 hr	—	No	Yes
Tobramycin	0.25	<5	Parenteral only	Renal	2 hr	2–4 days	Yes	No
Valproic acid	0.2	90	70–100	Hepatic	10–15 hr	—	No	Yes, use with caution

Therapeutic drug monitoring requires as much clinical skill as does titration of an oral anticoagulant dose by monitoring the prothrombin time. A basic assumption in this principle is that free drug at the *active site* is in equilibrium with total drug in plasma or serum. This has been shown probably to be true for many drugs. Furthermore, for these drugs, optimum therapeutic effects and minimal toxicity is observed when the serum drug concentration lies within an empirically determined therapeutic plasma concentration range. However, there is overlap between the therapeutic and subtherapeutic serum drug concentrations. Therefore, therapeutic drug monitoring should be considered as an aid to, not a substitute for, careful clinical observation in the management of drug therapy.

The purpose of this section is to provide some guidelines to follow for therapeutic drug monitoring and some of the salient features of the drugs being monitored. Table VI contains a list of drugs commonly monitored and the serum concentrations thought to represent the therapeutic range.

Interpretation of plasma drug concentrations clearly requires a broad knowledge of clinical pharmacokinetics. Recently, several sources of pharmacokinetic data have become available.

An appendix of pharmacokinetic data, developed by Benet and Sheiner,[15] is available. Included are excellent compilations of availability, urinary excretion, protein binding, clearance, volume of distribution, half-life and therapeutic and toxic concentrations for most of the currently used drugs. Data are accompanied by references so that the original work can be documented.

The newsletter, *Perspectives in Clinical Pharmacy*,[16] provides timely discussions of popular topics in clinical pharmacokinetics.

Another useful reference is by Gerson.[17] Included are chapters on the major drug classes with detailed discussions of the commonly used drugs.

The pharmacokinetics of abused substances are covered by Barnett and Chiang.[18]

Table VII provides important pharmacokinetic information for commonly monitored drugs. A sound knowledge of the clinical pharmacokinetics of each drug, a critical use of plasma drug concentrations as described above and a thorough clinical evaluation of the patient will provide the data required for the development of rational drug therapy.

References

1. Benet LZ: In McMahon FG, ed: *Principles and Techniques of Human Research and Therapeutics:* Futura Publishing, Mt Kisco NY, pp 9 to 23, 1974.
2. White PF: *Anesthiology 59:* 294, 1983.
3. Melander A *et al*: *Clin Pharmacol Therap 22:* 108, 1979.
4. Welling PG: In Bridges JW, Chasseavel LF, eds: *Progress in Drug Metabolism 4,* Wiley, New York, 1980.
5. Welling PG: In Benet LZ *et al*, eds: *Pharmacokinetic Basis for Drug Treatment,* Raven, New York, 131–163, 1984.
6. Ensminger WD *et al*: *Cancer Res 38:* 3784, 1978.
7. Klotz U *et al*: *J Clin Invest 55:* 347, 1975.
8. Gugler R *et al*: *Ibid:* 1182, 1975.
9. Spector AA *et al*: *Ann NY Acad Sci 226L:* 247, 1973.
10. Hori R *et al*: *Clin Pharmacol Therap 34:* 792, 1983.
11. Dettli L: *Clin Pharmacokinet 1:* 126, 1976.
12. Roberts RK *et al*: *Gastroenterology 75:* 479, 1978.
13. Kraus JW *et al*: *Clin Pharmacol Therap 24:* 411, 1978.
14. Shull HJ *et al*: *Ann Intern Med 84:* 420, 1976.
15. Benet LZ, Sheiner LB: In Gilman *et al*, eds: *The Pharmacological Basis of Therapeutics,* 7th ed, Macmillan, New York, 1663–1733, 1985.
16. *Perspectives in Clinical Pharmacy,* Elsevier, New York.
17. Gerson B: *Essentials of Therapeutic Drug Monitoring,* Igaku-Shoin, New York, 1983.
18. Barnett G, Chiang CN: *Pharmacokinetics and Pharmacodynamics of Psychoactive Drugs,* Biomedical Publications, Foster City CA.

CHAPTER 38

Topical Drugs

Stewart C Harvey, PhD

Professor of Pharmacology
School of Medicine, University of Utah
Salt Lake City, UT 84132

A large number of chemical agents may be applied to the skin and mucous membranes for their local effects. Many of these, such as antibiotics, antiseptics, corticosteroids, antineoplastics and local anesthetics, belong to distinct pharmacologic classes treated elsewhere in this text, and will not be discussed in this chapter. The remainder comprise a heterogeneous group of agents which, by exclusion, are mostly nonselective in action.

Those locally acting agents that have limited chemical and pharmacologic activity generally have a *physical* basis of action. Included in this group are protectives, adsorbents, demulcents, emollients and cleansing agents. The relative inertness of many of these substances renders them of value as vehicles and excipients. Consequently, many in this group are also pharmaceutical necessities and may be treated in Chapter 66.

Those locally acting agents that have general *chemical* reactivity include most astringents, irritants, rubefacients, vesicants, sclerosing agents, caustics, escharotics, many keratolytic (desquamating) agents and a miscellaneous group of dermatologics including hypopigmenting and antipruritic agents.

Although the skin and mucous membranes differ considerably in structure and function, they are similar in penetrability (to chemical agents) and in their response to certain physical and pharmacologic stimuli. Thus, many of the agents found in this chapter may be applied to both types of surfaces. Nevertheless, it is obvious that many agents, for which there is either contraindication or no rationale for their application to the mucous membranes, may be applied only to the skin.

In its broadest pharmacologic sense a protective is any agent that isolates the exposed surface (skin or mucous membrane) from harmful or annoying stimuli. In common practice only those substances that protect by mechanical or other physical means are considered to be protectives, although the surface action of adsorbents and demulcents cannot be divorced from their chemical properties. Protectives such as demulcents and emollients customarily are placed in separate categories; that practice is followed here.

The abridged category of protectives mainly comprises the dusting powders, adsorbents, mechanical protective agents and plasters.

Protectives and Adsorbents

Dusting Powders

Certain relatively indifferent (inert and insoluble) substances are used to cover and protect epithelial surfaces, ulcers and wounds. Usually these substances are subdivided very finely. They generally absorb moisture and, therefore, also act as cutaneous desiccants. The absorption of skin moisture decreases friction and also discourages certain bacterial growth.

The water-absorbent powders should not be administered to wet, raw surfaces because of the formation of cakes and adherent crusts. Starch and other carbohydrate powders not only may become doughy but they also may ferment. Consequently, such powders often contain an antiseptic. Most impalpable powders are absorptive, to some extent. Whether absorption of substances, other than water, contributes to the protection of the skin is uncertain; however, absorption of fatty acids and other constituents of perspiration, along with cutaneous drying, contributes to a deodorant action of the powders. It generally is held that the adsorptive capacity is important to the gastrointestinal protective action of chemically inert powders taken internally.

The chemically inert dusting powders are not entirely biologically inert, despite the name. When entrained in pores or wounds or left upon parietal surfaces, certain of the dusting powders, eg, talc, may cause irritation, granulomas, fibrosis or adhesions. Even without direct irritation or obstruction of the perspiration, dust can be troublesome.

Several of the dusting powders are incorporated into ointments, creams and lotions.

Bentonite—page 1305.
Boric Acid—page 1318.
Calcium Carbonate, Precipitated—page 776.
Talc—page 1327.
Titanium Dioxide—page 772.
Zinc Oxide—page 762.

Zinc Stearate

Octadecanoic acid, zinc salt

Zinc stearate [557-05-1]. A compound of zinc with a mixture of solid organic acids obtained from fats, and consists chiefly of variable proportions of zinc stearate and zinc palmitate. It contains the equivalent of 12.5–14.0% of ZnO(81.38).

Preparation—An aqueous solution of zinc sulfate is added to a sodium stearate solution, and the precipitate is washed with water until free of sulfate and dried.

Description—Fine, white, bulky powder, free from grittiness with a faint characteristic color; neutral to moistened litmus paper.

Solubility—Insoluble in water, alcohol or ether but is soluble in benzene.

Uses—In *water-repellent* ointments and as a *dusting powder* in dermatologic practice for its desiccating, astringent and *protective* effects. It has been removed from baby dusting powders, owing to accidental, fatal inhalations.

Mechanical and Chemical Protectives

Several materials may be administered to the skin to form an adherent, continuous coat which either may be flexible or semirigid, depending upon the substances and the manner in which they are applied. Such materials may serve three purposes: (1) to provide occlusive protection from the external environment, (2) to provide mechanical support and (3) to serve as vehicles for various medicaments.

The two principal classes of mechanical protectives are the collodions and plasters. Neither is used to much extent today. This is because there is increasing recognition of the beneficial effects of air in maintaining a normally balanced cutaneous bacterial flora of low pathogenicity. Also, the mechanical protectives may of themselves be somewhat irritating because of interference with normal water transport through the skin caused by certain oleaginous and resinous ingredients, especially in plasters. It also is recognized that rubber in adhesive plaster may induce eczema. The cerates may be employed similarly to the plasters. Bandages, dressings and casts also afford mechanical protection and support (see Chapter 105 for additional information). A brief discussion of plasters is included in Chapter 87.

A number of insoluble and relatively inert powders remain essentially unchanged chemically in the gastrointestinal tract. If the particles possess surface properties that favor their clinging to the gastrointestinal mucosa, and especially if they split up into tabular shapes, they offer mechanical protection against abrasion and may even offer slight protection against toxins and chemical irritants. Many such protectives also are adsorbents (charcoal, bismuth compounds, kaolin) or astringents (zinc and bismuth compounds). They are discussed under those categories.

Aluminum Hydroxide Gel—page 775.

Collodion

Contains not less than 5.0%, by weight, of pyroxylin.

Pyroxylin	40 g
Ether	750 mL
Alcohol	250 mL
To make about	1000 mL

Add the alcohol and the ether to the pyroxylin contained in a suitable container, and stopper the container well. Shake the mixture occasionally until the pyroxylin is dissolved.

Description—Clear, or slightly opalescent, viscous liquid; colorless, or slightly yellowish and has the odor of ether; specific gravity between 0.765 and 0.775.

Alcohol Content—22 to 26% of C_2H_5OH.

Uses—Chiefly to seal small wounds, for the preparation of medicated collodions and to protect nonaffected areas of the skin from topically applied irritants, corrosives, etc.

Caution—Collodion is highly flammable.

Flexible Collodion [Collodium Flexile]—See *RPS 16*, page 717. See also *Salicylic Acid Collodion* (page 768).

Absorbable Gelatin Film

Gelfilm (*Upjohn*)

A sterile, nonantigenic, water-insoluble, gelatin film obtained from a specially prepared gelatin-formaldehyde solution by drying on plates at constant temperature and humidity with subsequent sterilization by dry heat at 140° to 149°C for 12 hr.

Description—Light amber, transparent, pliable film that becomes rubbery when moistened.

Solubility—Insoluble in water; it assumes a rubbery consistency after being in water for a few minutes.

Uses—Both as a *mechanical protective* and as a *temporary supportive structure* and *replacement matrix* in *surgical repair of defects in membranes*, such as the dura mater and the pleura. When emplaced between damaged or operated structures, it prevents adhesions. When moistened, the film becomes pliable and plastic, so that it can be fitted to the appropriate surface. Absorption requires 1 to 6 months. It is also a component of stomadhesive, to be placed around an ostomy.

Dose—Applied in the form of sheets, previously soaked in isotonic sodium chloride solution and cut to the desired shape.

Dosage Forms—Film: 100 × 125 mm; Ophthalmic Film: 25 × 50 mm.

Zinc Gelatin

Zinc Gelatin Boot; Unna's Boot; Unna's Paste

Zinc Oxide	100 g
Gelatin	150 g
Glycerin	400 g
Purified Water	350 g
To make about	1000 g

Gradually add the gelatin to the cold purified water, with constant stirring, allow the mixture to stand for 10 min, and then heat on a steam bath until the gelatin dissolves. Add the zinc oxide, which previously has been rubbed to a smooth paste with the glycerin, and stir carefully until a smooth jelly result.

Uses—Melted and applied in the molten state between layers of bandage to act as a *protective* and to support varicosities and similar lesions of the lower limbs. After a period of about 2 weeks the dressing is removed by soaking with warm water.

Dose—*External*, as an occlusive boot.

Dosage Forms—Impregnated Gauze, in 10-yd lengths in following widths: 2¼, 2½, 3 and 4 in; impregnated with white or pink paste (the latter colored with a small amount of ferric oxide).

Kaolin—page 796.

Lanolin—page 1312.

Lanolin, Anhydrous—page 1311.

Mineral Oil—page 788.

Mineral Oil Emulsion—page 788.

Mineral Oil, Light—page 788.

Olive Oil—page 1309.

Peanut Oil—page 1303.

Petrolatum—page 788.

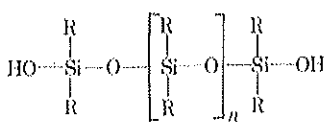
Other Mechanical and Chemical Protectives

Petrolatum Gauze [Petrolated Gauze]—Absorbent gauze saturated with white petrolatum. The weight of the petrolatum is 70–80% of the weight of the Gauze. It is sterile. Prepared by adding, under aseptic conditions, molten, sterile, white petrolatum to dry, sterile, absorbent gauze, previously cut to size, in the ratio of 60 g of petrolatum to each 20 g of gauze. *Uses:* A *protective* dressing; also as packing material for postoperative plugs, packs, rolls and tampons, and as a wick, drain or wrap-around for tubing. It is claimed that there is no danger of tissue maceration and that no growth of granulation tissue through it occurs.

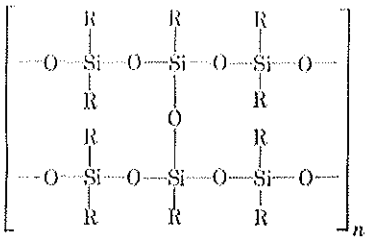
Dimethicone [Poly(dimethylsiloxane; poly[oxy(dimethylsilylene)] [9006-65-9] $(C_2H_6OSi)_n$]—A water-repellent silicone oil consisting essentially of dimethyl siloxane polymers of the 200 series of fluids (see *Silicones*, below). It is a water-white, viscous, oil-like liquid; immiscible with water or alcohol; miscible with chloroform or ether. *Uses:* Has skin-adherent and water-repellent properties. It is both a protective and an emollient, for which its FDA classification is Category 1. Applied to the skin, it forms a *protective* film that provides a barrier to ordinary soap and water and water-soluble irritants. The film may last several hours if the skin is exposed mainly to aqueous media. The film provides a less-effective barrier to synthetic detergents and lipid-soluble materials, such as organic solvents. It should not be applied except in contact dermatoses and dermatoses aggravated by substances that can be repelled by the silicone. It is useful in preventing irritation from ammonia produced by the urine of infants, but it may exacerbate preexisting irritation. The occlusive protection by the silicone is detrimental to inflamed, traumatized, abraded or excoriated skin and to lesions requiring free drainage. However, applied adjacent to such lesions, it offers protection against irritating discharges and maceration. It practically is harmless, and does not sensitize skin but it does cause temporary irritation to the eyes. It may be incorporated into ointments, creams and gels. *Dose:* Apply uniformly with rubbing 3 or 4 times for the first day or two, then twice daily. *Dosage Forms:* Aerosol, Cream and Ointment: 20 and 30%. All concentrations from 1 to 30% are approved.

Silicones (*Polyorganosiloxanes*)—These are organosilicon polymers containing chains of alternating oxygen and silicon atoms with substituent organic groups, frequently methyl or phenyl, attached to each silicon atom.

Preparation: These polymers may be prepared synthetically by condensing alkylated or arylated *silanols*. Disubstituted *silanediols* $[R_2Si(OH)_2]$ form linear polymers having the general formula:



Cross-linked polymers result from condensation of mixtures of substituted silanediols and monosubstituted *silanetriols* [$RSi(OH)_3$], represented by the following partial formula where R is a hydrocarbon radical:



One method of preparation involves interaction of silicon tetrachloride with appropriate Grignard reagents to yield alkylated or arylated dichlorosilanes. After hydrolysis to the corresponding substituted silanols, dehydration procedures are used to effect condensation polymerization. The overall reaction, as it involves a disubstituted silanediol, may be represented as:

$$SiCl_4 \xrightarrow{RMgX} R_2SiCl_2 \xrightarrow{HOH} R_2Si(OH)_2 \xrightarrow{-HOH} HO[Si(R)_2O]_nH$$

Silicon tetrachloride — **Disubstituted dichlorosilane** — **Disubstituted silanediol** — **Silicone linear polymer**

Properties: Silicones with a wide range of properties may be produced by varying the substituent R and the degree of cross-linking. Physically, silicones vary from mobile liquids through viscous liquids and semisolids to solids. Viscosities range from 0.65 to 1,000,000 centistokes. In general, they display high- and low-temperature stability. They are odorless, tasteless, relatively inert chemically and physiologically, water-repellent and possess antifoam characteristics. Unmodified silicones are generally insoluble in water; because of this the liquids often are termed *silicone oils;* however, a water-soluble sodium salt of a simple silicone, chemically *sodium methyl siliconate* [$CH_3Si(OH)_2ONa$], has been marketed.

Uses: Preparations containing silicones have various dermatological uses (see *Dimethicone*) and are used as ingredients of bases for ointments and liniments. In the form of inhalation sprays, silicone preparations have been employed in the treatment of pulmonary edema involving frothing of fluid in the upper respiratory tract. They also are used orally as antiflatulent or gastric defoaming agents (see *Simethicone*, page 799). A silicone *bouncing putty* has found acceptance for use as a physical agent in treating conditions requiring finger exercise. The water-repellent properties of the silicones have found considerable use in a great variety of applications where complete drainage of aqueous fluids from surfaces is desirable.

Silicones virtually are nonirritating; consequently, silicone rubbers are used in various indwelling catheters, tubes, etc, and in some types of prostheses. Liquid silicones are used also to fill in hypoplastic body areas for cosmetic purposes, although they tend to relocate because of flow under gravity and motion.

In addition to uses involving antifoaming, water-repellent and nonirritating characteristics, silicones also are employed to prevent sticking of one object to another and then are referred to commonly as release agents. Examples of such employment include release of rubber and plastics from molds, food from metal, ice from the wings of aircraft and capsules and tablets from molds and dyes in which they are fabricated.

Silicone rubbers are used to encapsulate steroid hormones and other drugs intended for chronic use, in order to retard absorption and effect a repository action lasting in some instances for as long as 1 yr. Continuing developments in this field offer interesting possibilities.

Zinc Carbonate [CO_3Zn(125.38)]—White rhombohedroids. Soluble 10 ppm in water at 15°; soluble in dilute acids, alkalies or solutions of ammonium salts. *Uses:* Both for its lubricity and as a drying agent. As a skin protectant it falls into FDA Category I. It is included in commercial topical burn and sunburn products and extemporary protectants. *Dose:* 0.2 to 2%.

Demulcents

Demulcents are protective agents that are employed primarily to alleviate irritation (*demulcere*—to smooth down), particularly of mucous membranes or abraded tissues. They also often are applied to the skin. They generally are applied to the surface in viscid sticky preparations that cover the area readily. The local action of chemical, mechanical or bacterial irritants, thereby, is diminished, and pain, reflexes, spasm or catarrh are attenuated. They also prevent drying of the affected surface. The demulcents may be applied to the skin in the form of lotions, cataplasms or wet dressing, to the gastrointestinal tract in the form of demulcent liquors or enemas and to the throat in the form of pastilles, lozenges or gargles. Demulcents also are included in artificial tears and in wetting agents for contact lenses. When demulcents are applied as solid material (as in lozenges or powders), the liquid is provided by secreted or exuded fluids. Demulcents frequently are medicated. In such instances the demulcent may be an adjuvant, a corrective or a pharmaceutical necessity. Many of the demulcents are also laxatives (page 783) and are used as such, or they are used with laxatives or antacids for their demulcent and lubricating action.

A variety of chemical substances possess demulcent properties. Among these are the alginates, mucilages, gums, dextrins, starches, certain sugars and polymeric polyhydric glycols. Mucus, in itself, is a natural demulcent. Certain silicates that form silicic acid on exposure to air or gastric juice and glycerin, although it is of low molecular weight and has relatively low binding power, frequently are placed among the demulcents. Also the colloidal hydrous oxides, hydroxides and basic salts of several metals are claimed to be demulcent, but acceptable clinical proof of the claim has not been provided.

The hydrophilic colloidal properties of most of the demulcents make them valuable emulsifiers and suspending agents in water-soluble ointments and suspensions. They also retard the absorption of many injections and, thus, may be employed in sundry depot preparations. Many of the demulcents mask the flavor of medicaments by means of at least three physical phenomena: (1) they apparently coat the taste receptors and render them less sensitive, (2) they incorporate many organic solutes into micelles and, thereby, diminish the free concentration of such solutes and (3) they coat the surfaces of many particles in suspension. Because of the adhesiveness of the demulcents, they are employed widely as binding agents in tablets, lozenges and similar dosage forms. Consequently, certain demulcents will be discussed in Chapter 66.

Acacia—page 1304.

Benzoin

Gum Benjamin; Benzoe

The balsamic resin obtained from *Styrax benzoin* Dryander or *Styrax parallelo neurus* Perkins, known in commerce as Sumatra Benzoin, or from *Styrax tonkinensis* (Pierre) Craib ex Hartwich, or other species of the Section *Anthostyrax* of the genus *Styrax*, known in commerce as Siam Benzoin (Fam *Styraceae*).

Sumatra benzoin yields not less than 75.0% of alcohol-soluble extractive, and Siam benzoin yields not less than 90.0% of alcohol-soluble extractive.

Constituents—Siam benzoin contains about 68% of crystalline *coniferyl benzoate* [$C_{17}H_{16}O_4$]; up to 10% of an amorphous form of this compound is also present. Some *coniferyl alcohol* (*m-methoxy-p-hydroxycinnamyl alcohol*, mp 73–74°) occurs in the free state as well. Other compounds that have been isolated are *benzoic acid* 11.7%, *d-siaresinolic acid* 6%, *cinnamyl benzoate* 2.3% and *vanillin* 0.3%.

Sumatra benzoin has been reported to contain benzoic and cinnamic acid esters of the alcohol *benzoresinol* and probably also of coniferyl alcohol, free *benzoic* and *cinnamic acids*, *styrene*, 2 to 3% of *cinnamyl cinnamate* (also called *styracin*), 1% of *phenylpropyl cinnamate*, 1% of *vanillin*, a trace of *benzaldehyde*, a little *benzyl cinnamate* and the alcohol *d-sumaresinol* [$C_{30}H_{48}O_4$].

Description—*Sumatra Benzoin:* Blocks or lumps of varying size made up of compacted tears, with a reddish brown, reddish gray or grayish brown resinous mass. *Siam Benzoin:* Compressed pebble-like tears of varying size and shape. Both varieties are yellowish to rusty

brown externally and milky white on fracture; hard and brittle at ordinary temperatures but softened by heat; aromatic and balsamic odor; aromatic and slightly acrid taste.

Uses—A *protective* application for irritations of the skin. When mixed with glycerin and water, the tincture may be applied locally for *cutaneous ulcers, bedsores, cracked nipples* and *fissures* of the lips and anus. For throat and bronchial inflammation, the tincture may be administered on sugar. The tincture and compound tincture sometimes are used in boiling water as steam inhalants for their *expectorant* and *soothing action* in acute laryngitis and croup. In combination with zinc oxide, it is used in baby ointments.

Dose—*Topical*, as a 10% tincture or compound tincture (below).

Compound Benzoin Tincture [Balsamum Equitis Sancti Victoris, Balsamum Commendatoris, Balsamum Catholicum, Balsamum Traumaticum, Balsamum Vulnerarium, Balsamum Persicum, Balsamum Succium, Balsamum Friari, Balsamum Vervaini, Guttae Nader, Guttae Jesuitarium, Tinctura Balsamica, Balsam of the Holy Victorious Knight, Commander's Balsam, Friar's Balsam, Turlington's Drops, Persian Balsam, Swedish Balsam, Vervain Balsam, Turlington's Balsam of Life, Balsam de Maitha, Ward's Balsam, Jerusalem Balsam, Saint Victor's Balsam, Wade's Drops, Wound Elixir and Balsamic Tincture]—*Preparation:* With benzoin (in moderately coarse powder, 100 g), aloe (in moderately coarse powder, 20 g), storax (80 g) and tolu balsam (40 g), prepare a tincture (1000 mL) by Process M (page 1543), using alcohol as the menstruum. *Alcohol Content:* 74 to 80% of C_2H_5OH. *Uses:* Especially valuable in acute *laryngitis*, also in croup, when added to hot water and the vapor inhaled. By adding a teaspoonful of the tincture to boiling water in an inhaler, and inhaling the vapor, very effective results may be obtained. See Chapter 104. Also administered, on sugar, for throat and bronchial inflammation and as a local application, when mixed with glycerin and water, for *ulcers, bedsores, cracked nipples* and *fissures* of the lips and anus. *Dose: Topical*, as required; *inhalation*, 1% in very hot water.

Carbomer Methylcellulose—page 1306.
Gelatin—page 1306.
Glycerin—page 931.
Glycerin Suppositories—page 785.
Glycyrrhiza—page 1295.
Hydroxypropyl Cellulose—page 1306.
Hydroxypropyl Methylcellulose—page 1306.
Hydroxyethyl Cellulose—page 1306.

Hydroxypropyl Methylcellulose Ophthalmic Solution

A sterile solution of hydroxypropyl methylcellulose, of a grade containing 19.0–30.0% methoxy and 4.0–12.0% hydroxypropoxy groups; may contain antimicrobial, buffering and stabilizing agents.

Uses—A wetting solution for contact lenses. Its demulcent action decreases the irritant effect of the lens on the cornea. It also imparts viscous properties to the wetting solution, which assists the lens in staying in place. The demulcent effect also finds application in ophthalmic decongestants. "Artificial tear" formulations containing this drug may be used when lacrimation is inadequate. A 2.5% solution is used in gonioscopes.

Dose—*Topical, to the conjunctiva*, **1 drop** of 0.3 to 1% solution 3 or 4 times a day.

Dosage Forms—0.3, 0.5 and 1% solutions.

Methylcellulose—page 1306.

Methylcellulose Ophthalmic Solution

A sterile solution of methylcellulose; may contain antimicrobial, buffering and stabilizing agents.

Uses—For the same purposes, and in the same manner, as *Hydroxypropyl Methylcellulose Ophthalmic Solution*, above.

Dosage Forms—0.25, 0.5 and 1%.

Pectin—page 796.
Polyvinyl Alcohol—page 1307.

Polyvinyl Alcohol Ophthalmic Solution

VasoClear A (*Cooper Vision*)

A sterile solution of polyvinyl alcohol, which may contain antimicrobial, buffering and stabilizing agents and other demulcent substances.

[9002-89-5] (Polyvinyl alcohol).

Preparation—By partial hydrolysis (*ca* 90%) of polyvinyl acetate.

Description—A white powder which is a linear polymer, $—(CH_2—CHOH)_n—$, where the value of n is between 500 and 5000; pH (1 in 25 aqueous solution) between 5.0 and 8.0.

Solubility—Soluble in water; insoluble in organic solvents.

Uses—A wetting solution for contact lenses. The polyvinyl alcohol has a demulcent action that helps protect the eye from irritation by the contact lens. It is also used in "artificial tears" employed when there is insufficient lacrimation. It is applied to the conjunctiva, **1 or 2 drops**, 3 or 4 times a day or as needed.

Dosage Forms—1, 1.4, 2, 3, and 4% solutions.

Emollients

Emollients are bland, fatty or oleaginous substances which may be applied locally, particularly to the skin, and also to mucous membranes or abraded tissues. Water-soluble irritants, air and airborne bacteria are excluded by an emollient layer. The skin also is rendered softer (*emollier*—to soften) and more pliable through penetration of the emollient into the surface layers, through the slight congestion induced by rubbing and massage upon application and especially through mechanical interference with both sensible and insensible water loss.

Emollients have certain disadvantages. It now is recognized that retention of perspiration below the emollient and exclusion of air render conditions favorable to the growth of anaerobic bacteria. Furthermore, the rubbing during application aids in the spreading of cutaneous bacteria. Consequently, the use of emollients to cover burns and abrasions is diminishing. The liquid emollients may be used for mild catharsis (page 783) and for protection against gastrointestinal corrosives; however, castor oil is hydrolyzed in the gut to the irritating ricinoleic acid and, hence, is employed as an emollient only externally. Orally administered liquid emollients may be aspirated into the trachea and lungs, especially in infants and in the debilitated, and, thus, induce "oil aspiration pneumonia." This condition also may be induced by emollient nose drops.

The chief use of emollient substances is to provide vehicles for lipid-soluble drugs (as in ointments and liniments), hence, many of them are described among the pharmaceutical necessities (Chapter 66). It is widely, but incorrectly, held that such vehicles facilitate the transport through the skin of their active ingredients. On the contrary, when the oil:water partition coefficient is greater than 1.0, the penetration is retarded and the emollient vehicle prolongs the action of the active ingredient. Emollient substances also are employed commonly in both cleansing and antiphlogistic creams and lotions. Compound ointment bases, creams and other medicated applications are treated elsewhere in this book (Chapter 86). Only the simple emollients and important compounded ointments that are used frequently for their emollient actions are listed below.

Castor Oil—page 785.
Castor Oil, Sulfated—page 1311.
Cocoa Butter—page 1611.
Coconut Oil—page 1317.
Cold Cream—page 1312.
Corn Oil—page 1303.

Cottonseed Oil—page 1303.
Ointment, Hydrophilic—page 1312.
Rose Water Ointment—page 1315.
Sesame Oil—page 1303.
Theobroma Oil—page 1320.
White Ointment—page 1309.
Yellow Ointment—page 1309.

Other Emollients

Myristyl Alcohol [Tetradecyl Alcohol |112-72-1| $CH_3(CH_2)_{12}CH_2OH$ (214.38)]—White crystalline alcohol; specific gravity 0.824; melts at 30°. Insoluble in water; soluble in ether; slightly soluble in alcohol. Obtained by reduction of fatty acid esters. *Use:* Emollient in cold creams.

Shark Liver Oil—The oil extracted from the livers of the *soupfin shark, Galeorhinus zyopterus or Hypoprion brevirostris*, both of which are rich in vitamins A and D. *Uses:* An emollient and protectant, the FDA classification of which is Category I. It is used in burn and sunburn ointments. *Dose:* Usually 3%.

Astringents and Antiperspirants

Astringents are locally applied protein precipitants which have such a low cell penetrability that the action essentially is limited to the cell surface and the interstitial spaces. The permeability of the cell membrane is reduced, but the cells remain viable. The astringent action is accompanied by contraction and wrinkling of the tissue and by blanching. The cement substance of the capillary endothelium and the basement membrane is hardened, so that pathological transcapillary movement of plasma protein is inhibited and local edema, inflammation and exudation, thereby, are reduced. Mucus or other secretions also may be reduced, so that the affected area becomes drier.

Astringents are used therapeutically to arrest hemorrhage by coagulating the blood (*styptic* action, page 816) and to check diarrhea, reduce inflammation of mucous membranes, promote healing, toughen the skin or decrease sweating. The *antiperspirant* effect is the result both of the closure of the sweat ducts by protein precipitation to form a plug and peritubular irritation that promotes an increase in inward pressure on the tubule. Astringents also possess some *deodorant* properties by virtue of interaction with odorous fatty acids liberated or produced by action of bacteria on lipids in sweat, and by an action suppressing bacterial growth, partly because of a decrease in pH.

Many astringents are irritants or caustics in moderate to high concentrations. Consequently, strict attention must be paid to the appropriate concentration. Most astringents are also antiseptics, hence, many of them are discussed in Chapter 62.

The principal astringents are (1) the salts of the cations aluminum, zinc, manganese, iron or bismuth, (2) certain other salts that contain these metals (such as permanganates) and (3) tannins, or related polyphenolic compounds. Acids, alcohols, phenols and other substances that precipitate proteins may be astringent in the appropriate amount or concentration; however, such substances generally are not employed for their astringent effects, because they readily penetrate cells and promote tissue damage. Strongly hypertonic solutions dry the affected tissues and, thus often, but wrongly, are called astringents, unless protein precipitation also occurs.

Alcohol—page 1314.

Alum

Sulfuric acid, aluminum potassium salt (2:1:1), dodecahydrate; Sulfuric acid, aluminum ammonium salt (2:1:1), dodecahydrate; Alumen; Alumen Purificatum; Purified Alum

Aluminum ammonium sulfate (1:1:2) dodecahydrate |7784-26-1|; anhydrous |7784-25-0| (237.14); or aluminum potassium sulfate (1:1:2) dodecahydrate |7784-24-9|; anhydrous |10043-67-1| (258.19).

The label of the container must indicate whether the salt is ammonium alum $[AlNH_4(SO_4)_2.12H_2O = 453.32]$ or potassium alum $[AlK(SO_4)_2.12H_2O = 474.38]$.

Preparation—Prepared from the mineral *bauxite* (a hydrated aluminum oxide) and sulfuric acid, with the addition of ammonium or potassium sulfate for the respective alums. Ammonium alum is prevalent on the market because of its lower cost.

Description—Large, colorless crystals, crystalline fragments or a white powder; odorless and has a sweetish, strongly astringent taste; solutions are acid to litmus.

Solubility—1 g ammonium alum is soluble in 7 mL water, and 1 g potassium alum is soluble in 7.5 mL water; both are soluble in about 0.3 mL boiling water, but they are insoluble in alcohol; alum is freely but slowly soluble in glycerin.

Incompatibilities—When alum is dispensed in powders with *phenol, salicylates* or *tannic acid*, gray or green colors may be developed due to traces of iron in the alum. A partial liberation of its water of crystallization permits it to act as an acid toward *sodium bicarbonate*, thus liberating carbon dioxide. Ammonia is liberated simultaneously from ammonium alum. *Alkali hydroxides and carbonates, borax* or *lime water* precipitate aluminum hydroxide from solutions of alum. The alums possess the incompatibilities of the water-soluble sulfates.

Uses—A powerful *astringent* in acidic solutions. It is slightly antiseptic, probably due to bacteriostasis through liberation of acid on hydrolysis. It sometimes is used as a local *styptic*, and frequently is employed in making astringent lotions and douches. It is used especially by athletes to toughen the skin. As an astringent it is used in concentrations of 0.5 to 5%. Some vulvovaginal cleansing and deodorant preparations contain alum.

Styptic pencils are made by fusing potassium alum, usually with the addition of some potassium nitrate, and pouring into suitable molds.

Caution—Do not confuse *styptic* pencils with *caustic* pencils (page 767); the latter contain *silver nitrate*.

Dose—*Topical*, as a **0.5 to 5%** solution.

Aluminum Acetate Topical Solution

Acetic acid, aluminum salt; Liquor Burowii; Burow's Solution

$Al(OCOCH_3)_3$

Yields, from each 100 mL, 1.20–1.45 g of aluminum oxide $[Al_2O_3 = 101.96]$, and 4.24 to 5.12 g of acetic acid $[C_2H_4O_2 = 60.05]$, corresponding to 4.8 to 5.8 g of aluminum acetate |139-12-8| $C_6H_9AlO_6$ (204.12). It may be stabilized by the addition of not more than 0.6% of boric acid.

Caution—This solution should not be confused with Aluminum Subacetate Topical Solution which is a stronger preparation.

Note—Dispense only clear Aluminum Acetate Solution.

Description—Clear, colorless liquid having a faint acetous odor, and a sweetish, astringent taste; specific gravity about 1.022; pH 3.6 to 4.4.

Uses—As an astringent dressing or as an astringent mouth wash and gargle. Aluminum acetate is included in preparations to treat athlete's foot, dermatidides, diaper rash, dry skin, poison ivy poisoning and inflammation of the external ear.

Dose—*Topical*, to the skin, as a wet dressing containing a **1:10 to 1:40 dilution** of the solution.

Aluminum Chloride

|7784-13-6| $AlCl_3.6H_2O$ (241.43); *anhydrous* |7446-70-0| (133.34).

Preparation—By heating aluminum in chlorine gas, then dissolving the product in water and crystallizing, or by dissolving freshly precipitated aluminum hydroxide in hydrochloric acid and concentrating to permit crystallization.

Description—White or yellowish white, crystalline powder; deliquescent; sweet, astringent taste; solutions are acid to litmus.

Solubility—1 g in about 0.9 mL water or 4 mL alcohol; soluble in glycerin.

Uses—Extensively employed on the skin as an astringent and anhidrotic; it is included in some proprietary preparations formu-

lated for this purpose. It is used especially in the treatment of soggy athlete's foot, to promote drying and, hence, to enhance the efficacy of specific antifungal drugs. For ordinary antiperspirant use the basic salt *aluminum chlorohydroxide*, $Al_2Cl(OH)_m$, is preferable as it is less irritating and causes less deterioration of clothing than does this drug. It may have a special use in the treatment of *hyperhidrosis of the palms, soles or axillae*, for which a 20% solution in absolute ethanol is used. In the presence of water, it hydrolyzes to aluminum chlorohydroxide and hydrochloric acid, which can cause irritation, especially in fissures, discomfort and also deterioration of clothing. Concentrations below 15% cause a low incidence of irritation. Consequently, it is essential that the area to be treated is completely dry before application. To protect bedclothes, the treated area is sometimes covered with plastic wrap, but such occlusion of the axillae may result in boils or furuncles. It should not be applied to the axillae immediately after shaving or used where the skin is irritated or broken. Concentrations above 15% are used as caustics.

Dose—*Topical*, to the skin, as 6.25 to 30% solution. The 20% alcoholic solution may be applied on 2 successive days and twice a week thereafter, except that it may be applied twice a day for athlete's foot.

Aluminum Chlorohydrates

The hydrate of aluminum chloride hydroxide [1327-41-9] $Al_2Cl(OH)_5$].

Uses—Mainly employed in antiperspirant products, for which they have been rated safe and effective in concentrations of 25% (as anhydride) or less. Since solutions or suspensions are less acidic than those of aluminum chloride, they cause a lower incidence of irritation to the skin.

Dose—*Topical*, to the axilla, as a 2.5 to 25% cake, ointment, solution or suspension.

Aluminum Sulfate

Sulfuric acid, aluminum salt (3:2), hydrate; Cake Alum; Patent Alum; Pearl Alum; Pickle Alum; "Papermaker's Alum"

Aluminum sulfate (2:3) hydrate [17927-65-0] $Al_2(SO_4)_3.xH_2O$; anhydrous [10043-01-3] (342.14).

Preparation—By reacting freshly precipitated aluminum hydroxide with an appropriate quantity of sulfuric acid. The resulting solution is evaporated and allowed to crystallize.

Description—White crystalline powder, shining plates or crystalline fragments; stable in air; odorless and has a sweet, mildly astringent taste; aqueous solution (1 in 20) is acid and has a pH not less than 2.9.

Solubility—1 g in about 1 mL water; insoluble in alcohol.

Uses—A powerful *astringent*, acting much like alum. It is used widely as a *local antiperspirant* and is the effective ingredient in some commercial antiperspirant products. Solutions usually are buffered with sodium aluminum lactate to make them less irritating. It is used for water purification in the "alum flocculation" process. It is a *pharmaceutical necessity* for *Aluminum Subacetate Solution*.

Dose—*Topical*, to the skin, as an 8% solution.

Bismuth Subcarbonate—page 799.

Bismuth Subnitrate—page 775.

Calamine

Iron oxide (Fe_2O_3), mixt. with zinc oxide; Prepared Calamine; Lapis Calaminaria; Artificial Calamine

Calamine [8011-96-9]; contains, after ignition, not less than 98.0% ZnO (81.38).

Preparation—By thoroughly mixing zinc oxide with sufficient ferric oxide (usually 0.5 to 1%) to obtain a product of the desired color.

It originally was obtained by roasting a native zinc carbonate, then known as *calamine*, hence, the name. This name also is applied by mineralogists to a native form of zinc silicate, which is not suitable for making medicinal calamine.

Description—Pink powder, all of which passes through a No 100 standard mesh sieve. It is odorless and almost tasteless.

Solubility—Insoluble in water; dissolves almost completely in mineral acids.

Uses—Similar to those of zinc oxide, being employed chiefly as an *astringent* and in *protective* and soothing ointments and lotions for *sunburn, ivy poisoning*, etc. It often is prescribed by dermatologists to give opacity and a flesh-like color to lotions or ointments.

Dose—*Topical*, to the skin, in various concentrations in lotions and ointments.

Calamine Lotion [Lotio Calaminae]—*Preparation:* Dilute bentonite magma (250 mL) with an equal volume of calcium hydroxide solution. Mix calamine (80 g) and zinc oxide (80 g) intimately with glycerin (20 mL) and about 100 mL of the diluted magma, triturating until a smooth, uniform paste is formed. Gradually incorporate the remainder of the diluted magma. Finally add calcium hydroxide solution (qs) to make 1000 mL, and shake well. If a more viscous consistency in the Lotion is desired, the quantity of bentonite magma may be increased to not more than 400 mL. *Note: Shake thoroughly before dispensing.*

Phenolated Calamine Lotion [Lotio Calaminae Composita; Compound Calamine Lotion]—*Preparation:* Mix liquefied phenol (10 mL) and calamine lotion (990 mL) to make 1000 mL. Commercial preparations also contain 8.4% isopropyl alcohol and have various other modifications. See *Calamine*. *Note: Shake thoroughly before dispensing.*

Glutaral—page 1165.

Potassium Permanganate—page 1173.

Resorcinol—RPS-16, page 1107.

Silver Nitrate—page 766.

White Lotion

Lotio Alba; Lotio Sulfurata

Zinc Sulfate	40 g
Sulfurated Potash	40 g
Purified Water, a sufficient quantity, To make	1000 mL

Dissolve zinc sulfate and sulfurated potash separately, each in 450 mL purified water, and filter each solution. Add slowly the sulfurated potash solution to the zinc sulfate solution with constant stirring. Then add the required amount of purified water, and mix.

Note—Prepare freshly and shake thoroughly before dispensing. For further discussion see *Sulfurated Potash* (page 1327).

Uses—An *astringent*, *protective* and mild antimicrobial preparation. The astringency is attributable to the zinc ion. The thiosulfates and polysulfides in it exert antibacterial and antifungal actions (see *Sodium Thiosulfate*, RPS-16, page 1176). White lotion is used in the treatment of acne vulgaris.

Dose—*Topical*, to the skin, as required.

Zinc Oxide

Flowers of Zinc; Zinc White; Pompholyx; Nihil Album; Lana Philosophica

Zinc oxide [1314-13-2] ZnO (81.38).

Preparation—By heating zinc carbonate at a low red heat until the carbon dioxide and water are expelled.

Description—Very fine, odorless, amorphous, white or yellowish white powder, free from gritty particles; gradually absorbs carbon dioxide from the air; when strongly heated it assumes a yellow color which disappears on cooling; its suspension in water is practically neutral.

Solubility—Insoluble in water or alcohol; soluble in dilute acids, solutions of the alkali hydroxides or ammonium carbonate solution.

Incompatibilities—Reacts slowly with fatty acids in *oils* and *fats* to produce lumpy masses of zinc oleate, stearate, etc. *Vanishing creams* tend to dry out and crumble. Whenever permissible, it is advisable to levigate it to a smooth paste with a little mineral oil before incorporation into an ointment.

Uses—Has a mild *astringent*, *protective* and *antiseptic* action. In the form of its various official ointments and pastes it is employed widely in the treatment of dry skin and such skin disorders and infections as *acne vulgaris, prickly heat, insect stings and bites, ivy poisoning, diaper rash, dandruff, seborrhea, eczema, impetigo, ringworm, psoriasis, varicose ulcers* and *pruritus*. It is contained in some sunscreens. It is included in some vulvovaginal deodorant preparations and in preparations for the treatment of hemorrhoids.

It also is used in dental cements and temporary fillings. It is the essential ingredient in *Calamine* (page 762).

Dose—*Topical*, as a 5 to 25% cream, lotion, ointment, paste, baby powder or rectal suppository.

Dosage Forms—Ointment: 20%; Paste: 25%. In numerous combinations: 2 to 15%.

Zinc Pyrithione—page 1173.

Zinc Sulfate—page 1170.

Zinc Undecylenate—page 1237.

Other Astringents and Antiperspirants

Aluminum Zirconium Chlorhydrate—*Uses:* Mainly in antiperspirant products. Because of the propensity of the zirconium to elicit allergic reactions and sarcoid-like granulomas, the compound is not included in aerosols, because of possible pulmonary complications if inhaled. *Dose:* To the axilla, in a concentration not to exceed 20% (as anhydride).

Tannic Acid [Gallotannic Acid; Tannin; Digallic Acid] [1401-55-4]—A tannin usually obtained from nutgalls, the excrescences produced on the young twigs of *Quercus infectoria* Olivier and allied species of *Quercus* Linné (Fam *Fagaceae*). Yellowish white to light brown amorphous powder, glistening scales or spongy masses; usually odorless with a strong astringent taste; gradually darkens on exposure to air and light. 1 g dissolves in about 0.35 mL water or 1 mL warm glycerin; very soluble in alcohol; practically insoluble in chloroform or ether. *Incompatibilities:* Solutions gradually darken on exposure to air and light through oxidation of phenolic groups to quinoid structures. It is incompatible with most enzymes, gums, salts of many metals and many other substances.

Uses: On an open sore or denuded surface, it forms a film of protein tannate that acts as a mechanical protective which excludes external irritants and infectives and, thus, provides some relief from pain. However, it is not antibacterial and not only does not inhibit the growth of bacteria entrained beneath the film but actually may create favorable conditions for the growth of certain anaerobes. For this reason, and also the fact that it is absorbed sufficiently from large denuded areas to cause liver damage, it is no longer used in the treatment of burns and should not be used on any large lesion. Nevertheless, it is incorporated in 8 to 10% concentration in several products to treat ivy or oak poisoning. As a 7% gel it is used on cold sores, fever blisters and cankers. It is included in 2.16% concentration in a hemorrhoidal preparation and in 4% concentration in a keratolytic product for removing corns, calluses and warts, these concentrations probably being too low to contribute significantly to the supposed efficacies. In 25% solution it is used to reduce inflammation and harden skin around ingrown toenails, thus increasing comfort and making nail-cutting easier.

Its content in tea accounts for the use of strong tea as an internal antidote, presumably for the dual purpose of precipitating toxic alkaloids and hardening the surface of the gastrointestinal mucosa and its mucous layer.

Zinc Caprylate [Zinc octanoate [557-09-5] $C_{16}H_{30}O_4Zn$ (351.79)]—Lustrous scales. Sparingly soluble in boiling water; moderately soluble in boiling alcohol. *Uses:* In the treatment of athlete's foot. The astringency of the zinc decreases inflammation and wetness. The caprylate has a weak antifungal action. *Dose:* As a 5% ointment.

Zinc Chloride [Zinc chloride [7646-85-7] $ZnCl_2$ (136.29)]—Prepared by reacting metallic zinc or zinc oxide with hydrochloric acid and evaporating the solution to dryness. White, or nearly white, odorless, crystalline powder, or as porcelain-like masses, or in moulded pencils; very deliquescent; aqueous solution (1 in 10) is acid to litmus. 1 g dissolves in 0.5 mL water, about 1.5 mL alcohol or about 2 mL glycerin; solution in water or alcohol is usually slightly turbid, but the turbidity disappears on addition of a small quantity of HCl. *Incompatibilities:* Soluble zinc salts are precipitated as zinc hydroxide by alkali hydroxides, including ammonium hydroxide; the precipitate is soluble in an excess of either the fixed or the ammonium hydroxide. *Carbonates, phosphates, oxalates, arsenates,* and *tannin* cause precipitation. The precipitation with sodium borate can be prevented by addition of an amount of glycerin equal in weight to the sodium borate. In weak aqueous solutions, it has a tendency to form the insoluble basic salt by hydrolysis and about one-half its weight of ammonium chloride has been used for the purpose of stabilization. It is very *deliquescent*. It has the incompatibilities of chlorides, being precipitated by *silver* and *lead salts*. *Uses:* In high concentrations it is caustic and has been used as a caustic agent to treat corns, calluses and warts. In the low concentrations in which it is marketed it is astringent and mildly antibacterial and probably does not contribute to keratolysis. Although it is used in mouthwashes, the contact time is too short, and only an astringent and not an antibacterial action results. *Dose: Topical*, to the teeth, as a 10% solution; to skin and mucous membranes for astringency and antimicrobial actions, as a 0.1 to 2% solution.

Zinc Ricinoleate [Zinc [*R*-(*Z*)]-12-hydroxy-9-octadecenoate $(C_{18}H_{33}O_3)_2Zn$ (660.24)]—Only as a deodorant for ostomies.

Zirconium Oxide [Zirconium Dioxide; Zirconic Anhydride, Zirconia; [1314-23-4] ZrO_2 (123.22)]—White powder or crystals. Insoluble in water; soluble in acids. *Uses:* Has weak astringent and adsorptive activity, for which it is employed in topical preparations for treating rhus dermatitis (ivy and oak poisoning). However, it is not only poorly effective for this purpose but it also can cause allergic reactions that may give rise to sarcoid-like granulomas. Consequently, its use should be condemned. Zirconium salts also are subject to the same criticisms.

Irritants, Rubefacients and Vesicants

The *irritants* are drugs that act locally on the skin and mucous membranes to induce hyperemia, inflammation and, when the action is severe, vesication. Agents that induce only hyperemia are known as *rubefacients*. Rubefaction is accompanied by a feeling of comfort, warmth and, sometimes, itching and hyperesthesia. Appropriately low concentrations of directly applied or inhaled vapors of volatile aromatic irritants, such as camphor or menthol, induce a sensation of coolness rather than warmth. When the irritation is more severe, plasma escapes from the damaged capillaries and forms blisters (vesicles). Agents that induce blisters are known as *vesicants*. Most rubefacients also are vesicants in higher concentrations. Certain irritants may be relatively selective for various tissues or cell types, so that hypersecretion of the surface, seborrheic abscesses, paresthesia or other effects may be noted in the absence of appreciable hyperemia.

Irritants have been used empirically for many centuries, probably even prehistorically. They may be employed for counterirritation, the mechanism of which is poorly understood. A moderate to severe pain may be obscured by a milder pain arising from areas of irritation appropriately placed to induce reflex stimulation of certain organs or systems, especially respiratory. Sensory and visible effects of irritation sometimes give the patient assurance that he is receiving effective medication. Taken internally, many irritants exert either an emetic or laxative action. Irritant laxatives are listed on page 783. A few irritants, especially cantharides, on absorption into the blood stream, irritate the urogenital tract and, consequently, have been dangerously employed as *aphrodisiacs*. Certain irritants also possess a healing action on wounds, possibly the result of local stimulation. Many condiments are irritants. In high concentrations, many irritants are corrosive.

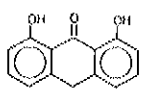
Alcohol—page 1314.

Alcohol, Rubbing—page 1164.

Ammonia Spirit, Aromatic—RPS-17, page 15.

Anthralin

1,8,9-Anthracenetriol; Dithranol;
Dioxyanthranol; Cignolin; Anthra-Derm (*Dermik*); Lasan (*Stiefel*)



1,8-Dihydroxyanthranol [480-22-8] $C_{14}H_{10}O_3$ (226.23).

Preparation—Anthraquinone is sulfonated to the 1,8-disulfonic acid, which is isolated from the reaction mixture and then heated with a calcium hydroxide-calcium chloride mixture to form 1,8-dihydroxy-9,10-anthraquinone, which is reduced with tin and HCl to anthralin.

Description—Yellowish brown, crystalline powder; odorless and tasteless; melts between 175° and 181°.

Solubility—Insoluble in water; slightly soluble in alcohol; soluble in chloroform; slightly soluble in ether.

Uses—Although long considered to be an irritant, its principal therapeutic action is the reduction of epidermal DNA synthesis and mitotic activity. It is used in the treatment of *psoriasis, alopecia areata, eczema* and other *chronic dermatoses*. It usually is used in

combination with ultraviolet light and a daily coal tar "bath." To avoid harmful irritation, medicaments containing it should not be used on the face, scalp, genitalia or intertriginous skin areas; they should not be applied to blistered, raw or oozing areas of the skin, and should be kept from the eyes, since they may cause severe conjunctivitis, keratitis or corneal opacity. Renal irritation, casts and albuminuria may result when the drug is absorbed systemically. The hands should be washed immediately after applying medication. A reversible slight discoloration of the skin may occur.

Dose—*Topical*, to the skin, as a 0.1 to 1% cream or ointment, once a day with cream and once or twice a day with ointment. The concentration should be low initially and increased only as necessary.

Dosage Forms—Cream: 0.1, 0.2, 0.25, 0.4, 0.5 and 1%; Ointment: 0.1, 0.25, 0.4, 0.5, 1 and 2%.

Benzoin Tincture, Compound—page 760.

Coal Tar

Pix Carbonis; Prepared Coal Tar BP; Pix Lithanthracis; Gas Tar

The tar obtained as a by-product during the destructive distillation of bituminous coal.

Description—Nearly black, viscous liquid, heavier than water, with a characteristic naphthalene-like odor and a sharp burning taste; on ignition it burns with a reddish, luminous and very sooty flame, leaving not more than 2% of residue.

Solubility—Only slightly soluble in water, to which it imparts its characteristic odor and taste and a faintly alkaline reaction; partially dissolved by alcohol, acetone, methanol, solvent hexane, carbon disulfide, chloroform or ether; to the extent of about 95% by benzene, and entirely by nitrobenzene with the exception of a small amount of suspended matter.

Uses—A *local irritant* used in the treatment of *chronic skin diseases*. Like anthralin, its primary action is to decrease the epidermal synthesis of DNA and, hence, to suppress hyperplasia. Occasionally, it may cause rash, burning sensation or other manifestations of excessive irritation or sensitization. Since photosensitization may occur, the treated area should be protected from sunlight. It should be kept away from the eyes and from raw, weeping or blistered surfaces. Temporary discoloration of the skin may occur.

Dose—*Topical*, to the skin: *cleansing bar*, 2% once or twice a day; *cream*, 1.6 to 5%, 2 or 3 times a day; *gel*, 5 to 7.5% once or twice a day; *lotion*, 2 to 5%, 2 to 4 times a day; *ointment*, 1 to 5%, 2 or 3 times a day; *paste*, 5% once or twice a day; *shampoo*, 0.5 to 10% twice a week; *solution*, 2.5 to 20% straight or diluted 1:3 with water 1 to 3 times a day; *suspension*, 7.5 to 33.3% diluted in lukewarm water at intervals directed by the physician.

Dosage Forms—Cleansing Bar: 2%; Cream: 1.6 and 5%; Gel: 5 and 7.5%; Lotion: 2 and 5%; Ointment: 1 and 5%; Paste: 5%; Shampoo: 0.5, 1, 2, 3, 4.3, 5, 9 and 10%; Topical Solution: 2.5, 5 and 20%; Topical Suspension: 7.5, 30 and 33.3%.

Green Soap—RPS-17, page 786.

Green Soap Tincture—RPS-17, page 766.

Methyl Salicylate—page 1295.

Resorcinol—RPS-16, page 1107.

Resorcinol Ointment, Compound—RPS-16, page 1107.

Resorcinol Monoacetate—RPS-16, page 1107.

Storax—page 1326.

Tolu Balsam—page 1299.

Turpentine Oil, Rectified—RPS-16, page 808.

Other Irritants, Rubefacients and Vesicants

Camphor [Bicyclo [2.2.1] heptane-2-one, 1,7,7-trimethyl-, 2-Camphanone; 2-Bornanone [76-22-2] $C_{10}H_{16}O$ (152.24); Gum Camphor; Laurel Camphor]—A ketone obtained from *Cinnamomum camphora* (Linné) Nees et Ebermaier (Fam *Lauraceae*) (Natural Camphor) or produced synthetically (Synthetic Camphor). *Preparation:* Natural crude camphor may be obtained by steam distilling chips of the camphor tree; the crude camphor so obtained is purified, usually by sublimation. One method of producing synthetic camphor starts with *pinene* [$C_{10}H_{16}$], a hydrocarbon obtained from turpentine oil. The pinene is saturated with hydrogen chloride at 0° forming bornyl chloride [$C_{10}H_{17}Cl$]. On heating the bornyl chloride with sodium acetate and glacial acetic acid, it is converted into isobornyl acetate, which is subsequently hydrolyzed to isobornyl alcohol [$C_{10}H_{17}OH$] and oxidized with chromic acid to camphor. Synthetic camphor resembles natural camphor in most of its properties except that it is a racemic mixture and, therefore, lacks optical activity. When camphor is mixed in approximately molecular proportions with chloral hydrate, menthol, phenol or thymol, liquefaction ensues; such mixtures are known as *eutectic mixtures* (see page 176).

Description: Colorless or white crystals, granules or crystalline masses; or as colorless to white, translucent, tough masses; a penetrating, characteristic odor, a pungent, aromatic taste and is readily pulverizable in the presence of a little alcohol, ether or chloroform; specific gravity about 0.99; melts between 174° and 179° and slowly volatilizes at ordinary temperature and in steam. *Solubility:* 1 g in about 800 mL water, 1 mL alcohol, about 0.5 mL chloroform or 1 mL ether; freely soluble in carbon disulfide, solvent hexane or fixed and volatile oils. *Incompatibilities:* Forms a liquid or a soft mass when rubbed with *chloral hydrate*, *hydroquinone*, *menthol*, *phenol*, *phenyl salicylate*, *resorcinol*, *salicylic acid*, *thymol* or other substances. It is precipitated from its alcoholic solution by the addition of water. It is precipitated from camphor water by the addition of soluble salts.

Uses: Locally, weakly *analgesic*, mildly *analgesic* (*antipruritic*) and *rubefacient* when rubbed on the skin. The spirit is applied locally to allay itching caused by insect stings. It also is used as a counterirritant in humans for *inflamed joints*, *sprains* and *rheumatic* and other *inflammatory* conditions such as colds in the throat and chest. Although the patient may feel improved, the inflammation is not affected. However, reflexly induced local vasoconstriction may mediate a mild nasopharyngeal decongestant effect. When taken internally in small amounts it produces a feeling of warmth and comfort in the gastrointestinal tract, and, therefore, formerly was much used as a *carminative*. Systemically, it is a reflexly active *circulatory* and *respiratory stimulant*. However, its use as a stimulant is obsolete. It also possesses a slight *expectorant* action and is included in some cough-suppressant mixtures. Concentrations above 11% are not safe. Toxicity consists of nausea and vomiting, headache, feeling of warmth, confusion, delirium, convulsions, coma or respiratory arrest. Camphor is a pharmaceutical necessity for *Flexible Collodion* and *Camphorated Opium Tincture*. *Dose: Topical*, to the skin, rectum or throat, as a 0.1 to 3% lotion, cream, spray or ointment, or 10% tincture (spirit), no more than 3 to 4 times a day. For topical analgesia, concentrations of 0.1 to 3% are used; for counterirritation, 3 to 11%.

Cantharidin [(3aα,4β,7β,7aα-Hexahydro-3a,7a-dimethyl-4,7-epoxyisobenzofuran-1,3-dione[56-25-7] $C_{10}H_{12}O_4$ (186.21)]—The active principle of *Cantharides*. White platelets soluble 1 g in 40 mL acetone, 65 mL chloroform, 560 mL ether or 150 mL ethyl acetate; soluble in oils. *Uses:* Produces intradermal vesiculation. It is used to remove warts, particularly the periungual type. It is applied under an occlusive bandage. The vesicle eventually breaks, becomes encrusted and falls off in 1 to 2 weeks. *Dose:* Topical, to the wart, as a 0.7% solution.

Capsicum—The dried ripe fruit of *Capsicum frutescens* Linné, *Solonaceae*, which contains less than 1% of capsaicin [(*E*)-*N*-[4-Hydroxy-3-methoxyphenyl)methyl]-8-methyl-6-nonaneamide[404-86-4] $C_{18}H_{27}NO_3$ (305.40), which is the active ingredient. *Uses:* Its active ingredients are mildly irritant, causing erythemia and a feeling of warmth without vesication. Its preparations are used as counterirritants. *Dose:* The equivalent of 0.025 to 0.25% of capsicum applied to the skin no more than 3 or 4 times a day.

Ichthammol [Ammonium Ichthosulfonate; Sulfonated Bitumen; Ichtiol; Ichthymall (*Mallinckrodt*), Ichthyol (*Stiefel*) [8029-68-3]]—It is obtained by the destructive distillation of certain bituminous schists, sulfonating the distillate and neutralizing the product with ammonia.

It yields not less than 2.5% of NH_3 (ammonia) and not less than 10% of total sulfur.

Constituents: It belongs to a class of preparations containing, as essential constituents, salts or compounds of a mixture of acids designated by the group name *sulfoichthyolic acid*, formed by sulfonation of the oil obtained in the destructive distillation of certain bituminous shales. Sulfoichthyolic acid is characterized by a high sulfur content, the sulfur existing largely in the form of sulfonates, sulfones and sulfides. *Description* and *Solubility:* Reddish brown to brownish black, viscous fluid, with a strong, characteristic, empyreumatic odor. Miscible with water, glycerin fixed oils or fats; partially soluble in alcohol or ether. *Incompatibilities:* Becomes granular in the presence of *acids* or under the influence of *heat*. In solution, it is precipitated by acids and *acid salts* as a dark, sticky mass; *alkalies* liberate ammonia; many *metallic salts* cause precipitation. *Uses:* A mildly astringent irritant and local antibacterial agent with moderate emollient and demulcent properties. It is used alone or in combination with other antiseptics for the treatment of skin disorders such as *insect stings and bites*, *erysipelas*, *psoria-*

sis and *lupus erythematosus* and to produce healing in *chronic inflammations*. It also is used to treat *inflammation* and *boils* in the external ear canal. Medical opinion is divided as to whether this agent is useful. In higher concentrations, irritation is frequent and rashes may develop. It should be kept away from the eyes and other sensitive surfaces. It has been reported to cause hyperepithelialization, an action that would be counterproductive in the treatment of psoriasis. *Dose: Topical*, to the skin as a 10 or 20% ointment or external ear canal as a 10% ointment.

Juniper Tar [Cade Oil]—The empyreumatic volatile oil obtained from the woody portions of *Juniperus oxycedrus* Linné (Fam *Pinaceae*). Dark brown, clear, thick liquid, having a tarry odor and a faintly aromatic, bitter taste. Very slightly soluble in water; 1 volume dissolves in 9 volumes of alcohol; dissolves in 3 volumes of ether, leaving a slight, flocculent residue; miscible with chloroform. *Uses:* A mildly irritant oil that is employed as a topical antipruritic in several chronic dermatologic disorders, such as *psoriasis, atopic dermatitis, pruritus, eczema* and *seborrhea*. Since it is irritant to the conjunctiva and also may cause chemosis of the cornea, care should be taken to keep it out of the eyes. Systemic absorption may result in renal damage. *Dose: Topical*, as 1 to 5% ointment applied once a day; it also is used as a 4% shampoo or 34% bath.

Menthol [Cyclohexanol, 5-methyl-2-(1-methylethyl)-, *p*-Menthan-3-ol; Peppermint Camphor [1490-04-6] $C_{10}H_{20}O$ (156.27)]—An alcohol obtained from diverse mint oils or prepared synthetically. It may be levorotatory [(−)-Menthol] from natural or synthetic sources, or racemic [(±)-Menthol].

Preparation: It owes its odor chiefly to menthol, which is obtained from it by fractional distillation and allowing the proper fraction to crystallize, or by chromatographic processes. Among numerous methods of synthesis of an optically inactive menthol, the most popular involves the catalytic hydrogenation of thymol (obtained from natural sources or synthesized from *m*-cresol or cresylic acid). The difficulty in the synthesis of (−)-menthol arises from the fact that menthol contains three asymmetric carbon atoms, and there are thus eight stereoisomers, designated as (−)- and (+)-menthol, (−)- and (+)-isomenthol, (−)- and (+)-neomenthol, and (−)- and (+)-neoisomenthol. To obtain a product meeting USP requirements, it is necessary to separate (−)-menthol from its stereoisomers, for which purpose fractional crystallization, distillation under reduced pressure or esterification may be used. The other stereoisomers differ from the official (−)-menthol in physical properties and possibly to some extent in pharmacologic action.

Description: Colorless, hexagonal, usually needle-like crystals, or fused masses, or a crystalline powder, with a pleasant, peppermint-like odor; (−)-menthol melts between 41° and 44°; (±)-menthol congeals at 27° to 28°. *Solubility:* Very soluble in alcohol, chloroform or ether; freely soluble in glacial acetic acid, mineral oil or in fixed and volatile oils; slightly soluble in water. *Identification:* When mixed with about an equal weight of camphor, chloral hydrate, phenol or thymol, it forms a "eutectic" mixture liquefying at room temperature. *Incompatibilities:* Produces a liquid or soft mass when triturated with *camphor, phenol, chloral hydrate, resorcinol, thymol* or numerous other substances. *Labeling:* The label on the container indicates whether it is levorotatory or racemic.

Uses: In low concentrations, selectively stimulates the sensory nerve endings for cold and, hence, causes a sensation of coolness. Some local analgesic effects also accompany this effect. Higher concentrations not only stimulate sensory endings for heat and other pain, but also may cause some irritation. Consequently, there may first be a sensation of coolness, then a slight prickly and burning sensation. The local analgesia and sensation of coolness are employed in the treatment of insect bites and stings, itching (antipruretic effect), minor burns and sunburn, hemorrhoids, toothache, cankers, cold sores and sore throat. The local analgesic effect also is the probable basis of the antitussive use, although the value of the drug as an antitussive remains unproved. Care must be taken to avoid the inhalation of irritant concentrations. The contribution of a placebo effect to some of these effects cannot be discounted. It is incorporated into irritant products used to treat acne vulgaris, dandruff, seborrhea, calluses, corns, warts and athlete's foot and in vaginal preparations to lessen the sense of irritation. Whatever effects the rubbing of menthol-containing ointment on the chest possess to relieve pulmonary congestion in colds and allergy are attributable to counterirritation and placebo effects. It also is contained in counterirritants for the treatment of muscle aches. *Dose: Topical*, to the skin, as a 0.1 to 2% lotion or ointment; to the throat, as a 0.08 to 0.12% lozenge. *Inhalation*, 15 mL of 1% liquid or 10 mL of 2% ointment per quart of water, to be dispensed by steam inhalation.

Peruvian Balsam [Peru Balsam; Balsam of Peru; Indian Balsam; Black Balsam]—Obtained from *Myroxylon pereirae* (Royle) Klotzsch (Fam *Leguminosae*). Contains from 60 to 64% of a volatile oil termed *cinnamein* and from 20 to 28% of *resin*. Cinnamein is a mixture of compounds, among which the following have been identified: the esters *benzyl benzoate, benzyl cinnamate, cinnamyl cinnamate* (*styracin*) and the alcohol *peruviol* (considered by some to be identical with the sesquiterpene alcohol *nerolidol*, $C_{15}H_{26}O$) as ester, free *cinnamic acid;* about 0.05% of *vanillin;* and a trace of *coumarin*. The resin consists of benzoic and cinnamic acid. *Description* and *Solubility:* Dark brown, viscid liquid; transparent and appears reddish brown in thin layers; agreeable odor resembling vanilla, a bitter, acrid taste, with a persistent after-taste and free from stringiness or stickiness. It does not harden on exposure to air; specific gravity 1.150 to 1.170. Nearly insoluble in water, but soluble in alcohol, chloroform or glacial acetic acid, with not more than an opalescence; partly soluble in ether or solvent hexane. *Uses: A local irritant and vulnerary*. It once was used as a dressing to promote growth of epithelial cells in the treatment of *indolent ulcers, wounds* and certain *skin diseases*, eg, scabies. It presently is an ingredient in suppositories used in the treatment of hemorrhoids and anal pruritus. Allergic reactions to it occassionally occur. Ointments containing both this and sulfur present a problem in compounding, since the resinous part of the balsam tends to separate. This difficulty may be overcome by mixing the balsam with an equal amount of castor oil, prior to incorporating it into the base; or alternatively, by mixing it with solid petroxolin [An ointment vehicle (oxygenated petroleum) consisting of liquid paraffin, oleic acid and ammoniated alcohol]. *Dose: Topical, rectal*, 1.8 to 30 mg in suppositories.

Pine Tar [Pix Pini; Pix Liquida; Tar]—The product obtained by the destructive distillation of the wood of *Pinus palustris* Miller, or of other species of *Pinus* Linné (Fam *Pinaceae*). Usually obtained as a by-product in the manufacture of charcoal or acetic acid from wood. It is a complex mixture of phenolic bodies for the most part insoluble in water. Among these are *cresol, phlorol, guaiacol, pyrocatechol, caerulignol* and *pyrogallol* ethers. Traces of *phenol* and *cresols* also are present as well as hydrocarbons of the paraffin and benzene series. *Description* and *Solubility:* Very viscid, blackish brown liquid; translucent in thin layers, but becomes granular and opaque with age; has an empyreumatic, terebinthinate odor, a sharp, empyreumatic taste and is more dense than water; solution is acid to litmus. Miscible with alcohol, ether, chloroform, glacial acetic acid or with fixed and volatile oils; slightly soluble in water, the solution being pale yellowish to yellowish brown. *Uses:* Externally as a mild *irritant* and *local antibacterial* agent in chronic *skin diseases*, especially eczema and psoriasis. Its volatile constituents are claimed to be *expectorant* but their efficacy is unproven; its inhalations were formerly used for this purpose. *Dose: Topical*, as a 1.8 to 30% shampoo.

Sclerosing Agents

A number of irritant drugs are of sufficient activity to damage cells but are not so potent as to destroy large numbers of cells at the site of application. Such agents promote fibrosis and are used to strengthen supporting structures, close inguinal rings, etc. The intimal surface of blood vessels may break down under attack by such agents and thus initiate thrombosis, which may be an undesirable side effect. This action is the basis of the use of sclerosing agents in the reduction of varicose veins and hemorrhoids. Sclerosing agents generally are regarded as obsolete. They can be harmful when improperly used and sometimes even when used with caution.

Sclerosing Agents

Morrhuate Sodium Injection—A sterile solution of the sodium salts of the fatty acids of cod liver oil. It contains 50 mg of sodium morrhuate/mL. A suitable antimicrobial agent, not to exceed 0.5%, and ethyl or benzyl alcohol, not to exceed 3%, may be added. *Note: It may show a separation of solid matter on standing. Do not use the material if such solid does not dissolve completely upon warming.* Prepared by heating cod liver oil with alcoholic sodium hydroxide until completely saponified. After dilution with water the alcohol is removed by distillation. Dilute H_2SO_4 is then added to the aqueous solution, and the liberated organic acids are separated or preferably extracted with a suitable immiscible solvent such as ether. Just-sufficient aqueous NaOH then is added to neutralize the acids. About 20 mg of benzyl alcohol/mL of the Injection usually is added to lessen the pain of injection. *Uses:* Formerly, widely used as a *sclerosing* and *fibrosing agent* for obliterating *varicose veins*. Irritants of this type once were employed for closure of hernial rings, fibrosing of uncomplicated hemorrhoids, removal of condylomata acuminata and in other conditions where the ultimate objective was production of fibrous tissue. *Dose: Intravenous*, by special injection, 0.5 to 5 mL of a 5% injection to a localized area; *usual*, 1 mL. *Dosage Forms:* 5 and 30 mL.

Sodium Tetradecyl Sulfate [7-Ethyl-2-methyl-4-undecanol hydrogen sulfate sodium salt [139-88-8] $C_{14}H_{29}NaO_4S$ (316.43); STS; Sotradecol Sodium (*Elkins-Sinn*)]—One method of preparation reacts the corresponding alcohol with $ClSO_3H$ and neutralizes the resulting hydrogen sulfate ester with Na_2CO_3. Occurs as a white, waxy, odorless solid. Soluble in water, alcohol or ether. *Uses:* A *sclerosing agent* similar in action to sodium morrhuate. It formerly was used widely as a buffered solution in the *obliteration of varicose veins and internal hemorrhoids.* For such purposes, the solution is injected directly into the vein. Injection outside of the vein may cause sloughing. For this reason, the substance is not used to close inguinal rings. The principal untoward effect is pain immediately upon injection, although brief; mild anaphylactoid and idiosyncratic responses rarely occur. Because the substance is an anionic surface-active agent, it also is used as a wetting agent to promote spreading of certain topical antiseptics. *Dose:* By injection directly into the target vein, as a 1 or 3% solution, depending on the size of the vein. The volume then to be injected at any one site varies from 0.2 to 2.0 mL, depending on the concentration and the number of previous injections at the site, the larger volumes being given only after several previous injections. No more than 10 mL of the 3% solution or 6 mL of the 5% solution should be given at any one time. The interval between injections varies from 5 to 7 days. *Dosage Form:* Injection: 1 and 3% in 2-mL ampuls.

Caustics and Escharotics

Any topical agent that causes destruction of tissues at the site of application is a *caustic* (or corrosive).

Caustics may be used to induce desquamation of cornified epithelium ("keratolytic" action) and, therefore, are used to destroy warts, condylomata, keratoses, certain moles and hyperplastic tissues.

If the agent also precipitates the proteins of the cell and the inflammation exudate, there is formed a scab (or eschar), which later is organized into a scar; such an agent is an *escharotic* (or cauterizant). Most, but not all, caustics are also escharotic. Furthermore, certain caustics, especially the alkalies, redissolve precipitated proteins, partly by hydrolysis, so that no scab or only a soft scab forms; such agents penetrate deeply and generally are unsuitable for therapeutic use. Escharotics sometimes are employed to seal cutaneous and aphthous ulcers, wounds, etc. Since most escharotics are bactericidal, it formerly was thought that chemical cauterization effected sterilization; however, sterilization is not achieved always, especially by those agents which remain bound to the protein precipitate. The growth of certain bacteria even may be favored by the chemically induced necrosis and by the protection of the scab.

Acetic Acid, Glacial—page 1317.
Alum—page 761.
Aluminum Chloride—page 761.
Phenol—page 1323.

Podophyllum

Mandrake; May Apple

The dried rhizome and roots of *Podophyllum peltatum* Linné (Fam *Berberidaceae*); it yields not less than 5% of podophyllum resin.

Constituents—From 3 to 6% of resin along with up to 1% of quercetin and podophyllotoxin and peltatin glucosides. At least 16 different compounds have been isolated and characterized. The aglycone *podophyllotoxin* [$C_{22}H_{22}O_8$] is the lactone of 1-hydroxy-2-(hydroxymethyl)-6,7-methylenedioxy-4-(3′,4′,5′-trimethoxyphenyl)-1,2,3,4-tetrahydronaphthalene-3-carboxylic acid. Hydrolytic rupture of the lactone ring yields *podophyllic acid* [$C_{22}H_{24}O_8$], the 2,3-*trans* form of which is *podophyllinic acid* while the 2,3-*cis* form is *picropodophyllinic acid.*

Although podophyllotoxin has been demonstrated to possess marked caustic, cathartic and toxic properties, it is believed that not it, but an amorphous *resin*, called *podophylloresin*, is the chief cathartic principle of the drug. However, podophyllotoxin is safer and ultimately probably will replace the crude preparations.

Uses and Dose—See *Podophyllum Resin.*

Podophyllum Resin

Uses—Supersedes podophyllum (above). Certain glycosides and polynuclear lactones in the resin interact with tubulin and, thus, interfere with cell cycling and intracellular dynamics such as to cause the eventual death of affected cells. Applied topically, it is corrosive in the region of contact. It mainly is used in the treatment of *condyloma accuminatum* but also of *juvenile papilloma of the larynx, multiple superficial epitheliomatoses* (basal cell and squamous cell carcinomas), *precancerous keratoses* (seborrheic, actinic and radiation keratoses), *verrucae fibroids* and *calluses.* Some pain usually occurs at the site of application; if it is excessive, the drug should be removed with ethanol or isopropyl alcohol. Resin on adjacent normal tissues also should be removed. Pain may be avoided somewhat by treating only a small area of surface at any one time. *It especially is irritating to the eyes and mucous membranes.* Treatment of large surfaces also may result in excessive absorption and systemic effects, such as nausea and vomiting, tachycardia, shallow respiration, leukopenia, thrombocytopenia, renal damage, paralytic ileus, lethargy, stupor, psychotic confusional states and peripheral neuropathy, including flaccid paralysis. Systemic absorption is enhanced by occlusion. The drug is contraindicated in pregnancy and lactation.

Dose—*Topical, adults* and *children*, to the *skin*, *condyloma accuminata*, as a 25% solution, the resin to remain in place for 6 hr; application may be repeated weekly for up to 4 weeks, if necessary; *superficial epitheliomatoses* and *precancerous keratoses*, as a 25% solution once a day, to be continued until several days after a slough has occurred; *to laryngeal lesions, juvenile laryngeal papilloma*, as a 12.5% solution to the papilloma, initially once a day, but progressively longer intervals may be elected as the lesions shrink (medical authorities hold that short intervals are more effective); the 12.5% solution is to be extemporized by diluting the 25% solution in 95% ethanol.

Dosage Form—Topical Solution: 25%.

Salicylic Acid—page 768.
Silver Nitrate—page 766.

Silver Nitrate

Nitric acid silver(1+) salt; Argenti Nitras

Silver(1+) nitrate [7761-88-8] $AgNO_3$ (169.87).

Preparation—By the action of nitric acid on metallic silver.

Description—Colorless or white crystals; on exposure to light in the presence of organic matter, it becomes gray or grayish black; pH of solutions about 5.5.

Solubility—1 g in 0.4 mL water, 30 mL alcohol, about 250 mL acetone, slightly more than 0.1 mL boiling water or about 6.5 mL of boiling alcohol; slightly soluble in ether.

Incompatibilities—Easily reduced to metallic silver by most *reducing agents*, including *ferrous salts, arsenites, hypophosphites, tartrates, sugars, tannins, volatile oils* and other *organic substances.* In neutral or alkaline solutions, precipitated by *chlorides, bromides, iodides, borax, hydroxides, carbonates, phosphates, sulfates, arsenites* and *arsenates. Potassium permanganate, tannic acid* and *soluble citrates and sulfates* may cause a precipitate if sufficiently concentrated. In acid solution, only the *chloride, bromide* and *iodide* are insoluble. *Ammonia water* dissolves many of the insoluble silver salts through formation of the silver diammine complex, $Ag(NH_3)_2^+$.

Uses—Silver ions combine with proteins and cause denaturation and precipitation. As a result, silver ions have astringent, caustic, bactericidal and antiviral properties. In low concentrations, silver-denatured protein is confined to the interstitial spaces and the surface of denuded, weeping areas, so that only astringent and antimicrobial effects occur; with higher concentrations, cell membranes are disrupted, so that caustic effects result. The corroded site will become covered with a scab of silver protein precipitate.

It is used mainly in podiatry as a caustic to *destroy excessive granulation tissue*, such as *corns, calluses, granuloma pyogenicum* and *plantar warts*, to *reduce neurovascular helomas, remove papillomas* and *cauterize small nerve endings* and *blood vessels.* As an astringent, it is used to treat *impetigo vulgaris* and *pruritis* as well as *indolent ulcers, wounds* and *fissures.* It also is used as a *styptic*, especially in dentistry.

As an antiseptic, it mainly is employed prophylactically against *ophthalmia neonatorum.* It formerly was applied regularly to

burned surfaces because of its high efficacy against both staphylococci and pseudomonas. However, the precipitation of AgCl at the site of application and in dressing depletes plasma choride and can cause serious electrolyte disturbances; consequently, the drug seldom is used in burn therapy today. Refer to RPS-17, page 1165, for a discussion of its prior uses as an antiseptic.

Excessive corrosion at the target site and corrosion from inadvertent application or leakage away from the intended site can occur. Dental cones or pieces of toughened silver nitrate that are accidentally ingested can cause death. Elemental silver from the bioreduction of silver ion may reside permanently at the site of application and cause a bluish-to-black discoloration called argyria. Locally injected sodium thiosulfate sometimes can remove the silver. Nitrate ion absorbed from large, denuded surfaces can cause methemoglobinemia. Only concentrations 0.5% or below should be applied to raw wounds, fresh cuts or broken skin.

Dose—*Topical, antiseptic*, to the *conjunctiva*, **0.1 mL** of a **1%** solution; to the *burned skin* or *open lesion* (neither advised), **0.1** to **0.5%** solution as a wet dressing. *Astringent*, to the *affected skin*, as a **10%** solution for *impetigo vulgaris* and as a **10** or **25%** solution for *pruritis*. *Caustic*, to the *lesion* only, as a **10%** solution or ointment for *helomas* and to *cauterize small nerve endings* and *blood vessels*, as a **25** or **50%** solution for *plantar warts* and as a **50%** solution for *granulation tissue, granuloma pyogenica* and *papillomatous growths*.

Dosage Forms—Ointment: 10%; Topical Solution: 10, 25 and 50%. For *Toughened Silver Nitrate*, see RPS-17, page 784.

Other Caustics and Escharotics

Dichloroacetic Acid [Dichloroacetic acid $C_2H_2Cl_2O_2$ (128.95)]—Pungent liquid miscible with water, alcohol or ether. *Uses:* See *Trichloroacetic Acid*.

Nitric Acid—Contains 67–71% HNO_3. A fuming liquid, very caustic, with a characteristic, highly irritating odor; boils at 120°; specific gravity about 1.41. Miscible with water. *Uses:* As a cauterizing agent for the immediate sterilization of dangerously infected wounds, such as the bite from a rabid animal; it does not penetrate too deeply and forms a firm eschar.

Podophyllotoxin—[(5*R*,5a*R*,9*R*)-5,5a,6,8,8a,9-Hexahydro-9-hydroxy-5-(3,4,5-trimethoxyphenyl)furo[3′,4′:6,7]naphtho[2,3-*d*]-1,3-dioxol-6-one [518-28-5] $C_{22}H_{22}O_8$ (414.41)]—Found in the rhizomes of several species of plants, principally *Podophyllum peltatum* L. *Berberidaceae*, *P emodi* and *Juniperus virginiana* L, *Coniferae*. For the synthesis see *JACS 103:* 6208, 1981. Occurs as hydrated crystals; melts about 115° (dec) and about 184° after drying; a number of polymorphic forms exist. Very slightly soluble in water; soluble in alcohol, chloroform or acetone. *Uses:* Actions, uses and adverse effects are those of *Podophyllum Resin* (page 766), except that the therapeutic index is greater. It is several times more potent. It is an investigational drug. *Dose: Topical*, to the *skin*, *adults* and *children*, as a 0.5 to 1% solution twice a day for 3 days.

Potassium Hydroxide [Potassium hydroxide; Caustic Potash; Lye; Potash Lye [1310-58-3] contains not less than 85.0% of total alkali, calculated as KOH (56.11), including not more than 3.5% of K_2CO_3 (138.21)] *Caution—Exercise great care in handling, as it rapidly destroys tissues. Do not handle it with bare hands.* Prepared by electrolysis of a solution of potassium chloride in a diaphragm cell that does not allow liberated chlorine to react with it. It is prepared in the form of sticks, pellets, flakes or fused masses. Sticks or pellets are made by evaporating a solution of it to a fluid of oily consistency and then pouring the hot liquid into suitable molds in which it solidifies. *Description* and *Solubility:* White, or nearly white, fused masses, small pellets, flakes, sticks, and other forms; hard and brittle and shows a crystalline fracture; exposed to air it rapidly absorbs carbon dioxide and moisture, and deliquesces; melts at about 360–380°; when dissolved in water or alcohol, or when its solution is treated with an acid, much heat is generated; solutions, even when highly diluted, are strongly alkaline. 1 g dissolves in 1 mL water, 3 mL alcohol or 2.5 mL glycerin at 25°; very soluble in boiling alcohol. *Incompatibilities:* Bases react with *acids* to form salts, liberate alkaloids from aqueous solutions of *alkaloidal salts*, and promote various hydrolysis reactions such as the decomposition of *chloral hydrate* into chloroform and a formate or the breakdown of *salol* into phenol and a salicylate. Only the alkali hydroxides are appreciably soluble in water. Nearly all common *metals* will be precipitated as hydroxides when solutions of their salts are added to solutions of the alkali hydroxides. Certain hydroxides, however, notably those of aluminum, zinc, arsenic or lead, will dissolve in excess of sodium or potassium hydroxide. *Uses:* A *caustic*, principally in veterinary practice. The end of a stick of potassium hydroxide may be inserted into a section of rubber tubing, or wrapped several times with tin foil, to avoid cauterizing the fingers of the operator. It is used also as a *pharmaceutical necessity* in several pharmacopeial preparations.

Trichloroacetic Acid [Acetic acid, trichloro-, Trichloroacetic acid [76-03-9] $C_2HCl_3O_2$ (163.39)]—Usually made by oxidizing chloral hydrate with fuming nitric acid. Colorless, deliquescent crystals having a slight, characteristic odor; melts at about 58° and boils at 196°–197°. *Solubility:* 1 g in about 0.1 mL water; soluble in alcohol or ether. *Uses:* Precipitates proteins and used as a *caustic* on the skin or mucous membranes to destroy local lesions and for treatment of various dermatologic diseases. Its chief use is to destroy ordinary warts and juvenile flat warts. It is employed extensively as a precipitant of protein in the chemical analysis of body fluids and tissue extracts, as well as a decalcifier and fixative in microscopy. *Caution—Trichloroacetic Acid is highly corrosive to the skin. Dose: Topical*, to the skin, as a 15 to 100% *w/v* solution, carefully applied with a cotton-tipped applicator or glass rod. Concentrations above 50% are not recommended.

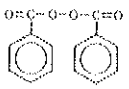
Zinc Chloride—page 763.

Keratolytics (Desquamating Agents)

The epidermis consists of layers of flat cells, called *stratified squamous epithelial cells*. They are bound together by desmosomes and penetrating tonofibrils, both of which largely consist of keratin. The outer layer of the epidermis, the cornified epithelium or stratum corneum, is made up of the collapsed ghosts of the squamous cells and, as such, is principally a tight network of keratin and lipoprotein. Certain fungi, especially the dermatophytes, utilize keratin and, therefore, reside in the stratum corneum in those places where the degree of hydration and the pH are sufficiently high. One way such mycoses may be suppressed is that of removal of the stratum corneum, a process that is called *desquamation*. Certain chemical substances, especially among phenols and sulfhydryl compounds, loosen the keratin and, thus, facilitate desquamation. These substances are called *keratolytics*. Aqueous maceration of the stratum corneum also favors desquamation. In addition to the treatment of epidermophytosis, keratolytics are used to thin hyperkeratotic areas. Most keratolytics are irritant. Irritants also can cause desquamation by causing damage to and swelling of the basal cells.

Benzoyl Peroxide

(Various Mfrs)



[94-36-0] $C_{14}H_{10}O_4$ (242.23); contains 65–82% of benzoyl peroxide; also contains about 26% of water for the purpose of reducing flammability and shock sensitivity.

Preparation—Benzoyl chloride is reacted with a cold solution of sodium peroxide.

Description—White, granular powder, having a characteristic odor; melts about 104°; *may explode with heat.*

Solubility—Sparingly soluble in water or alcohol; soluble in acetone, chloroform or ether.

Caution (For the drug entity—not the dosage forms)—It may explode at temperatures higher than 60° or cause fires in the presence of reducing substances. Store it in the original container, treated to reduce static charges. Do not transfer it to metal or glass containers fitted with friction tops. Do not return unused material to its original container, but destroy it by treatment with NaOH solution (1 in 10) until addition of a crystal of KI results in no release of free iodine.

Uses—Possesses mild antibacterial properties, especially against anaerobic bacteria. It is also mildly irritant, and it exerts moderate keratolytic and antiseborrheic actions. Its principal use is in the treatment of mild *acne vulgaris* (in which it is comedolytic) and *acne rosaceae*, but it also is used in the treatment of decubital and stasis ulcers.

It causes stinging or burning sensations for a brief time after application; with continued use these effects mostly disappear. After 1 or 2 weeks of use there may be a sudden excess dryness of the

skin and peeling. The drug must be kept away from the eyes, and from inflamed, denuded or highly sensitive skin, such as the circumoral areas, neck and skin of children. It should not be used in conjunction with harsh abrasive skin cleansers. It can cause contact dermatitis. It can bleach hair and fabrics.

Dose—*Topical*, to the *skin*, *adults* and *children 12 yr or older*, as a 5 or 10% cleansing bar 2 or 3 times a day, 5 to 10% cream or gel 1 or 2 times a day, 5 to 20% lotion 1 to 4 times a day, 5 or 10% cleansing lotion 1 or 2 times a day, 5% facial mask once a day, 10% soap 1 or 2 times a day or 10% stick 1 to 3 times a day. The 20% lotion is used only for the treatment of decubital and stasis ulcers.

Dosage Forms—Cleansing Bar: 5 and 10%; Cream: 5, 7 and 10%; Gel: 2.5, 5 and 10%; Lotion: 5, 5.5, 10 and 20%; Cleansing Lotion: 5 and 10%; Facial Mask: 5%; Stick: 10%.

Fluorouracil—page 1151.

Resorcinol—RPS-16, page 1107.

Resorcinol Ointment, Compound—RPS-16, page 1107.

Salicylic Acid

Benzoic acid, 2-hydroxy-, *o*-Hydroxybenzoic Acid

Salicylic acid [69-72-7] $C_7H_6O_3$ (138.12).

Preparation—Mostly by the Kolbe-Schmidt process in which CO_2 is reacted with sodium phenolate under pressure at about 130° to form sodium salicylate, followed by treatment with mineral acid.

Description—White, fine, needle-like crystals or as a fluffy, white, crystalline powder; the synthetic acid is white and odorless; sweetish, afterward acrid, taste; stable in the air; melts between 158° and 161°.

Solubility—1 g in 460 mL water, 3 mL alcohol, 45 mL chloroform, 3 mL ether, 135 mL benzene or about 15 mL boiling water.

Uses—Used *externally* on the skin, where it exerts a slight *antiseptic* action and considerable *keratolytic* action. The latter property makes it a beneficial agent in the local treatment of certain forms of *eczematoid dermatitis*. It also is included in products for the treatment of *psoriasis*, for which the FDA classification is Category 1. Tissue cells swell, soften and ultimately desquamate. Salicylic Acid Plaster often is used for this purpose. The drug is especially useful in the treatment of *tinea pedis* (athlete's foot) and *tinea capitis* (ringworm of the scalp), since the fungus grows and thrives in the stratum corneum. Keratolysis both removes the infected horny layer and aids in penetration by antifungal drugs. It is combined with benzoic acid in an ointment long known as Whitfield's Ointment. It also is combined commonly with zinc oxide, sulfur or sulfur and coal tar. It is incorporated into mixtures for the treatment of acne, dandruff and seborrhea, insect bites and stings and into soaps and vaginal douches, but efficacy remains to be established. In high concentrations it is *caustic* and may be used to remove *corns*, *calluses*, *warts* and other growths.

Collodions or solutions of 17% or higher and other forms above 25% concentration should not be employed if the patient has diabetes mellitus, peripheral vascular disease or inflammation or infection at the intended site of application. Continuous application of the drug to the skin can cause dermatitis. Systemic toxicity resulting from application to large areas of the skin has been reported. It is not employed internally as an analgesic because of its local irritating effect on the gastrointestinal tract.

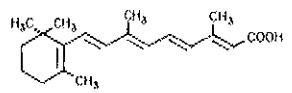
Dose—*Topical*, to the *skin*, *keratolytic*, as a 16.7 or 17% collodion once a day, 2.5 to 10% cream under occlusion once every 3 to 5 days, 2% foam once or twice a day, 5 or 6% gel under occlusion once a day, 1.8% lotion once or twice a day, 3 to 10% ointment once a day, 2 or 4% shampoo once or twice a week, 3.5% soap once a day or 17% solution once a day; *antipsoriatic*, as a 5 or 6% gel under occlusion or 3 to 10% ointment once a day; *antiseborrheic*, as a 1.8% lotion, 3 to 10% ointment or 2 or 4% shampoo once a day; *antiacne*, as a 2% foam once or twice a day, 5 or 6% gel under occlusion once a day, 3 to 6% ointment once a day or 3.5% soap once a day; *caustic*, as a 25% cream once every 3 to 5 days, 25 to 60% ointment under occlusion every 3 to 5 days, or 40% plaster once a day.

Dosage Forms—Flexible Collodion: 16.7 and 17%; Cream: 2.5, 10 and 25%; Gel: 5 and 6%; Lotion: 1.8%; Ointment: 25, 40 and 60% (3 to 10% ointments must be extemporized); Plaster: 40%; Shampoo: 2 and 4%; Soap: 3.5%; Topical Solution: 17%.

Sulfur, Precipitated—page 1247.

Tretinoin

Retinoic acid; Retin-A (*Ortho*)



all *trans*-Retinoic acid [302-79-4] $C_{20}H_{28}O_2$ (300.44).

Preparation—By oxidation of vitamin A aldehyde which may be obtained by oxidation of vitamin A. *Biochem J 90:* 569, 1964.

Description—Yellow to light-orange crystals or crystalline powder with the odor of ensilage; should be stored in cold and protected from light and air; melts between 176 and 181°.

Solubility—Insoluble in water; slightly soluble in alcohol; slightly soluble in chloroform; 1 g in 10 mL boiling benzene.

Uses—It is retinoic acid, or so-called *vitamin A acid*, which is formed when the aldehyde group of retinene (retinal) is oxidized to a carboxyl group. It is not known whether retinoic acid has a physiologic function, but some authorities consider it to be the form of vitamin A that acts in the skin. This view is supported by the fact that retinol and retinal have very little action on the skin but large systemic doses of vitamin A evoke prominent dermatologic changes.

Topically, it causes inflammation, thickening of the epidermis (acanthosis) and local intercellular edema, which leads to some separation of the epidermal cells. Follicular epithelial cells become less adhesive, the stratum corneum loosens and exfoliation may occur. High concentrations can cause vesiculation. These actions are used in the treatment of *acne vulgaris*. The loosened horny layer makes it easier for the comedo to rise up and discharge, and the inflammatory response mobilizes white cells which attack the bacteria in the follicle. In the early stages of treatment, the sudden surfacing of obscured preexisting comedones makes it appear that the acne has been exacerbated, but the new comedones do not coalesce into cysts or nodules and scarring does not occur. The exaggerated stage may last for as long as 6 weeks, after which improvement comes rapidly. Shortly after discontinuation of treatment, relapses readily occur. Deep cystic nodular acne (acne conglobata) or severe cases usually are not improved by the drug.

Various hyperkeratotic conditions are reported to respond to it, responses being sometimes exceptionally dramatic. *Solar* and *follicular keratosis*, *lamellar ichthyosis*, *keratosis palmaris* and *plantaris* and other hyperplastic dermatoses have been treated successfully with the drug. It also has been used in the treatment of some skin cancers. Recent reports indicate that it may somewhat rejuvenate sun-aged skin.

It is an antioxidant and free-radical scavenger. There is some evidence not only that topical applications may provide some protection from actinic and other radiation effects on the skin, including cancer, but that internally it may be protective against carcinogenesis from radiation and carcinogens. Systemically, it does not cause the toxic effects of large doses of vitamin A.

In concentrations of 0.05 to 0.1%, it causes a transient feeling of warmth or mild stinging, and erythema follows. Peeling of the skin may occur. Irritation and peeling are marked more when the concentration exceeds 0.1%. When peeling, crusting or blistering occurs, medication should be withheld until the skin recovers, or the concentration should be reduced. The drug should not be applied around the eyes, nose or angles of the mouth, because the mucosae are much more sensitive than the skin to the irritant effects. It also may cause severe irritation on eczematous skin. It should not be applied along with, or closely following, other irritants or keratolytic drugs. Exposure to sunlight should be avoided if possible. Both hypo- and hyperpigmentation have been reported, but the conditions appear to be reversible and temporary.

Dose—*Topical, usual*, to the skin, 0.01 to 0.1% once a day at bedtime.

Dosage Forms—Cream: 0.05 and 0.1%; Gel: 0.01 and 0.025%; Topical Solution: 0.05%.

Trichloroacetic Acid—page 767.

Urea—page 931.

Cleansing Preparations

The skin may be cleansed with detergents, solvents or abrasives, singly or in combination. Among the detergents, the soaps have enjoyed the greatest official status, more through custom than through special merit. The nonsoap detergents became important, not only as household hand cleansers, but also in dermatologic and surgical practice as well. However, because many nonsoap detergents do not decompose in sewage disposal plants, there has been a return to real soap. Some of the antiseptic "soaps" still contain synthetic detergents. Soap interferes with the action of many antiseptics, which is one reason synthetic detergents often are used in antiseptic cleansing preparations. However, synthetic detergents also interact with some antiseptics. Anionic nonsoap skin detergents rarely sensitize the skin and, thus, are prescribed when the user is allergic to soap.

Ordinary soaps tend to be alkaline, with pH ranging from 9.5 to 10.5. Superfatted soaps have a pH in the lower end of the range. Synthetic detergents usually have a pH of 7.5 or less. Neutral toilet bars contain synthetic detergents.

Shampoos are liquid soaps or detergents used to clean the hair and scalp. Both soaps and shampoos often are used as vehicles for dermatologic agents.

Many bar soaps contain either triclosan or triclocarban as antiseptics in concentrations which suppress bacterial production of body odors but which effectively are not antiseptic. A number of soaps and shampoos contain keratolytic and antiacne ingredients. Abrasive soaps contain particles of alumina, polyethylene or sodium tetraborate decahydrate.

It commonly, but erroneously, is believed that soap has an antiseptic action. The promotion of either soap or synthetic detergents alone for the control of acne is unwarranted; antiseptic substances must be added to the cleansing material or be used separately. Quantitative studies of the cutaneous flora before and after cleansing with soap or with other anionic detergents show a negligible antiseptic effect. However, the removal of loose epidermis lessens the likelihood that cutaneous bacteria will be transferred from the skin to other structures. Certain cationic detergents employed in dermatology are antiseptic. Detergents are treated under *Surface-Active Agents* (page 267).

The choice of organic solvents to cleanse the skin depends largely upon the nature of the material to be removed. In medical practice ethanol and isopropyl alcohol are the most frequently employed organic solvents. Cleansing creams act both as solvents and as detergents. Other soapless cleansers variously contain petrolatum, vegetable oils, lanolin, high-molecular-weight alcohols, various carbohydrate derivatives, oatmeal and other ingredients.

Alcohol—page 1314.
Alcohol, Rubbing—page 1164.
Benzalkonium Chloride—page 1164.
Green Soap—RPS-17, page 786.
Hexachlorophene Cleansing Emulsion—page 1166.
Isopropyl Rubbing Alcohol—page 1167.
Sodium Lauryl Sulfate—page 1307.

Miscellaneous Dermatologics

Gargles, nasal washes, douches, enemata, etc generally contain as basic ingredients substances described under other categories in this chapter. These preparations are described under *Aqueous Solutions*, page 1521.

Antiphlogistics include alcohol and several creams and lotions that cool the skin by evaporation. Many antiphlogistic preparations also contain an astringent and a local anesthetic or camphor or menthol.

Commonly employed *antipruritics* also depend largely upon local anesthetics and the soothing effect of cooling, although emollients or demulcents may be included, especially depending upon the etiology of the pruritus. The antipruritic properties of phenol preparations largely derive from superficial local anesthesia.

Vulnerary and *epithelizing* properties are attributed to numerous irritants and to several dyes; however, few reliable data exist to support most claims to vulnerary action.

Sunscreens contain aromatic compounds, like aminobenzoic acid, which efficiently absorb the harmful ultraviolet (UV) rays from the incident sunlight and transmit mainly the less harmful wavelengths, or titanium dioxide, which reflects sunlight from the surface of application. UV light in the spectral range of 290–320 nm causes suntan and sunburn; therefore, a sunscreen to prevent tan or burn should have a high molar absorptivity in this range. However, *photosensitization* (ie, the photoactivation of chemicals to make them toxic or allergenic) may occur with wavelengths as high as 500 nm; consequently, to protect recipients of certain drugs (tetracyclines, sulfonamides, erythromycin, promazine, chlorpromazine, promethazine, psoralens), sunscreens with a broader absorption spectrum are required. An adequate broad spectrum is usually achieved with combinations of sunscreens (eg, dioxybenzone and oxybenzone).

Melanizers are substances that promote the pigmentation of the skin. Most melanizers produce their effect by sensitizing the skin to UV light,* so that the effect is principally the same as if the subject had been exposed for a long time to the sun.

Skin bleaches, or *demelanizers*, mostly contain hydroquinone derivatives.

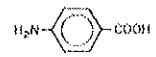
Hair bleaches generally contain peroxides.

There is a large variety of *depilatories* on the market. Many of them are sulfhydryl compounds, especially thioglycollates, which reduce the disulfide bonds of keratin, thus softening the hair to the point where it can be separated easily from the epidermis. Some of the same compounds are used in lower concentrations in hairwaving preparations. There is one drug, minoxidil, an antihypertensive drug, which can *increase hair growth* and *treat baldness*. Diazoxide probably will prove to have similar activity.

Antiperspirants have been included among the astringents.

Aminobenzoic Acid

Benzoic acid, 4-amino-, PABA



p-Aminobenzoic acid [150-13-0] $C_7H_7NO_2$ (137.14).

Preparation—*p*-Nitrotoluene is oxidized with permanganate to *p*-nitrobenzoic acid, and the nitro group is then reduced to amino with iron and hydrochloric acid.

Description—White or slightly yellow, odorless crystals or crystalline powder; melts between 186° and 189°; discolors on exposure to air or light.

* This action is termed a *photodynamic action*. The term has been used loosely to include all instances of enhanced sensitivity to light, but in strict definition it is confined to photosensitization in which the participation of oxygen is required. In the photodynamic process, light of wavelengths too long to be ordinarily effective may be used, so that the activating spectrum may be shifted toward longer wavelengths.

Solubility—Slightly soluble in water or chloroform; freely soluble in alcohol or solutions of alkali hydroxides and carbonates; sparingly soluble in ether.

Uses—A *sunscreen*. It absorbs UV light of wavelengths in the region of 260 to 313 nm; its molar absorptivity at 288.5 nm is 18,300. However, it does not absorb throughout the near UV range, so that drug-related photosensitivity and phototoxicity may not be prevented by it, but in combination with benzophenone it does protect against some drug-induced phototoxicities. Nevertheless, in the 260–313 nm range, it has the highest protection index of current sunscreen agents.

For animal species that do not use preformed folic acid, which contains the *p*-aminobenzoyl moiety, it is a B-vitamin. However, man does not use it, and its promotion in vitamin preparations preys on the ignorance of the consumer. It or its potassium salt is promoted as an agent that softens or regresses fibrotic tissue in Peyronie's disease, scleroderma, dermatomyositis, morphea and pemphigus. The claims for the antifibrotic actions are substantiated poorly, and the actions and uses are not mentioned in major works on pharmacology and therapeutics.

Topically, it is rarely allergenic to recipients but phototoxicity and photoallergenicity occur. Systemic side effects include nausea, anorexia, fever and rash.

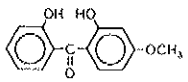
Dose—*Topical*, as a sunscreen, 4 to 15% in solutions, lotions, creams and lipsticks. *Oral, adults*, 12 g a day in 4 to 6 divided doses; *children*, 1 g/10 lb a day in divided doses, to be diluted and taken with food.

Dosage Forms—Capsules: 500 mg; Cream: 4% (may also contain sodium PABA); Gel: 5%; Lotion: 5%; Powder: 2, 100 and 453 g; Solution: 5%; Stick: 5% (may contain red petrolatum); Tablets: 30, 100 and 500 mg.

Cetyl Alcohol—page 1312.

Dioxybenzone

Methanone, (2-hydroxy-4-methoxyphenyl)(2-hydroxyphenyl)-, Spectra-Sorb UV 24 (*American Cyanamid*); Solaquin (*Elder*)



2,2′-Dihydroxy-4-methoxybenzophenone [131-53-3] $C_{14}H_{12}O_4$ (244.25).

Preparation—By a Friedel-Crafts reaction in which *o*-methoxybenzoyl chloride is added gradually to a mixture of 1,3-dimethoxybenzene, chlorobenzene and aluminum chloride. The reaction conditions are such that both methoxy groups ortho to the carbonyl bridge in the initial condensation product are demethylated. US Pat 2,853,521.

Description—Off-white to yellow powder; congeals not lower than 68°.

Solubility—Practically insoluble in water; freely soluble in alcohol or toluene.

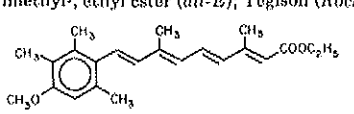
Uses—A *sunscreen* of intermediate molar absorptivity (11,950 at 282 nm), but it absorbs throughout the UV spectrum and, hence, affords protection not only against sunburn but also against the photodynamic, photosensitizing and phototoxic effects of drugs. At present, it is marketed in combination with the closely related *Oxybenzone* (page 771).

Dose—*Topical*, as a 3% lotion.

Dosage Forms—Dioxybenzone and Oxybenzone Cream: 3% of each ingredient.

Etretinate

2,4,6,8-Nonanetetraenoic acid, 9-(4-methoxy-2,3,6-trimethylphenyl)-3,7-dimethyl-, ethyl ester (*all-E*); Tegison (*Roche*)



[54350-48-0] $C_{23}H_{30}O_3$ (354.49).

Preparation—One scheme involves the Wittig condensation of diphenyl 2,3,6-trimethyl-4-methoxybenzylphosphonium chloride and 8-oxo-3,7-dimethyl-2,4,6-octane-trienoic acid (*all-trans*) in the presence of butylene oxide; *Experientia 34:* 1113, 1978.

Description—Crystalline solid melting about 104°.

Uses—Although not a topical drug, it is a retinoid closely related to tretinoin and is used only for its dermatologic actions; consequently, it is included in this chapter. It is used in the treatment of recalcitrant *psoriasis*, especially the severe pustular erythrodermic type. It decreases scaling, erythema and the thickness of lesions and causes epithelial and dermal cells to redifferentiate to normal cells. Sometimes, dramatic improvement occurs within 2 weeks and complete clearing in 1.5 to 4.5 mo. However, relapses are frequent once treatment is discontinued and sometimes even during chronic maintenance. It can be used alone or in low-dose combination with PUVA therapy. The mechanism of action is unknown, but it is undoubtedly like that of vitamin A. Activity resides in the acid metabolite.

Adverse effects occur in more than 75% of recipients. They include chapped lips, peeling of the palms, soles and fingertips, dryness of the mucous membranes, sore tongue, cheilitis, rhinorrhea, nosebleed, gingival bleeding, loss of hair, nail abnormalities, dry and irritated cornea, sclera and conjunctiva (50%), epidermal fragility, easy sunburning and other effects. Occasionally, pseudotumor cerebri, metastatic calcification of ligaments and tendons, and liver dysfunction or necrosis occur. In children and adolescents there may be premature closure of the epiphyses. Plasma cholesterol and triglycerides rise and high-density lipoprotein decreases. The drug is also teratogenic. Adverse effects are less with the low doses used with PUVA.

Absorption after oral administration is incomplete. It is increased by whole milk and other lipid-containing foods. There is a rapid metabolism during which it is deesterified to the acid metabolite. A much slower degradation and conjugation follows, the metabolites being secreted into bile and urine. Nearly all of the circulating drug is bound to plasma lipoproteins, but the active metabolite is bound to albumin. Ultimately, it is taken up into fat, where it may be found even as long as 2 yr after the last dose. The apparent elimination half-life is about 120 days. This persistence of drug in the body militates against the use of the drug in fertile women of child-bearing age, since the incidence of congenital defects is high even when conception occurs months after the drug is discontinued. The drug also is excreted into milk; effects in the nursing infant are not known.

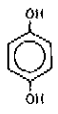
Dose—*Oral, adult, initially* **0.25 to 1.5 mg/kg** a day in divided doses, the dose depending upon the type and seriousness of the disorder; with erythrodermic psoriasis, the initial dose is 0.25 mg/kg a day, increased weekly with increments of 0.25 mg/kg a day until a response occurs; *maintenance*, **0.5 to 0.75 mg/kg** a day. Maintenance usually is not begun until after 8 to 16 weeks of treatment. The above doses are higher than those used concurrently with PUVA treatment.

Dosage Form—Capsules: 10 and 25 mg.

Hydrogen Peroxide Solution—page 1171.

Hydroquinone

1,4-Benzenediol; *p*-Dihydroxybenzene; Hydroquinol; Quinol; Eldoquin and Eldopaque (*Elder*)



Hydroquinone [123-31-9] $C_6H_6O_2$ (110.11).

Preparation—Various processes are employed. One involves reacting a sulfuric acid solution of aniline with manganese dioxide and reducing the resulting *p*-benzoquinone with sodium bisulfite.

Description—Fine, white needles; darkens on exposure to air; melts between 172 and 174°.

Solubility—1 g in about 17 mL water, 4 mL alcohol, 51 mL chloroform or 16.5 mL ether.

Uses—A *hypopigmenting* agent employed percutaneously to lighten localized areas of hyperpigmented skin, such as skin blem-

ishes, lentigo, melasma, chloasma, freckles, etc. Its action is temporary, so that it is necessary to repeat the application at frequent intervals. It is a mild irritant, and erythema or rash may develop, which requires discontinuation of the drug. It should not be used near the eyes or in open cuts. It is contraindicated in the presence of sunburn, miliaria or irritated skin. It is not to be used in children.

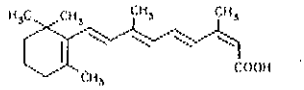
Dose—*Topical*, to the skin, *adults* and *children over 12 yr* as a 2 to 4% cream, gel, lotion or ointment to the affected area once or twice a day.

Dosage Forms—Cream: 2 and 4%; Gel: 4%; Lotion: 2%; Ointment: 2 and 4%.

Hydroxyurea—page 1158.

Isotretinoin

13-*cis*-Retinoic Acid; Accutane (*Roche*)



3,7-Dimethyl-9-(2,6,6-trimethyl-1-cyclohexen-1-yl)-2-*cis*-4-*trans*-6-*trans*-8-*trans*-nonatetraenoic acid [4759-48-2] $C_{20}H_{28}O_2$ (300.44). Differs from tretinoin (vitamin A) only in the configuration of the unsaturation at the α and β carbon atoms, which is *cis* rather than *trans*.

Uses—Although not a topical drug, it is a dermatologic agent and, hence, is described here. Its primary action is to decrease the production of sebum, which lends itself to the treatment of severe *modular* and *cystic acne* (acne conglobata). The size of the sebaceous gland is decreased and there is a change in the morphology and secretory capacity of the cells (dedifferentiation). Complete clearing of lesions is seen in about 90% of cases. A single course of treatment usually brings about long-lasting, sometimes permanent, remissions.

It also appears to diminish hyperkeratosis and has been reported to be effective in *rosacea*, gram-negative *folliculitis*, *lamellar ichthyosis*, *Darier's disease*, *pityriasis rubra pilaris* and *keratocanthoma*.

Adverse effects include facial dermatitis, fragile skin, thinning and drying of the hair, reversible cheilitis and dry skin, mouth, eyes and conjunctivitis in 25 to 80% of recipients. Peeling of the palms and soles and sensitivity to sunburn occur in about 5% of users. Urethral inflammation also occurs frequently. Joint pains and exacerbation of rheumatoid arthritis also has been reported to occur in about 16% of patients. Sedimentation rate, serum triglyceride concentration and serum levels of alanine and aspartate transaminases transiently occur in about 25% of users. In spite of the relatively high incidence of side effects, treatment rarely has to be discontinued.

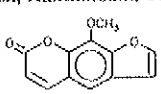
After oral administration, peak blood concentrations occur within 1 to 4 hr. The compound is oxidized to 4-hydroxy-13-*cis*-retinoic acid, which is then glucuronidated and is secreted into the bile. The elimination half-life is 11 to 39 (mean 20) hr. Isotretinoin should not be given during pregnancy or nursing.

Dose—*Oral, adult*, for *acne*, 1 to 2 **mg/kg** a day in 2 divided doses for 15 to 20 weeks. If the cyst count has not been reduced by more than 70%, a second course of treatment may be given after a wait of 2 months. Persons over 70 kg or who have severe chest and back involvement usually require doses at the high end of the range. For *severe rosacea* or gram-negative *folliculitis*, **0.25** to **0.5 mg/kg** twice a day. For *hyperkeratoses*, up to **4 mg/kg**.

Dosage Forms—Capsules: 10, 20 and 40 mg.

Methoxsalen

7*H*-Furo [3,2-*g*][1]benzopyran-7-one, 9-methoxy-, Ammoidin; 9-Methoxypsoralen; Xanthotoxin; Oxsoralen (*Elder*)



[298-81-7] $C_{12}H_8O_4$ (216.19).

Preparation—Occurs naturally in *Psorales coryfolia*, *Ammi majus*, *Ruta chalepensis* and various other plants. It may be synthesized by methods described in *JACS 79:* 3491, 1957, and in US Pat 2,889,337.

Description—White to cream-colored, odorless, fluffy, needle-like crystals; melts between 143° and 148°.

Solubility—Practically insoluble in cold water, sparingly soluble in boiling water; freely soluble in chloroform; soluble in boiling alcohol, acetone or acetic acid; soluble in aqueous alkalies with ring cleavage; reconstitution occurs on neutralization.

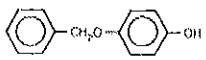
Uses—A psoralen melanizer. It increases the photodynamic pigmentation of skin; it does not induce pigmentation in the absence of UV light or melanocytes. It is used in the treatment of *vitiligo* and to *desensitize to sunlight*. Severe sunburning can occur with topical application; it is customary to protect the surrounding skin with a sunscreen. It also is used in PUVA treatment of *psoriasis*, *mycosis fungoides* and *cutaneous T-cell lymphoma;* in these, irradiation activates it to cross-link DNA. It may have value in the PUVA treatment of *alopecia areata*, *inflammatory dermatoses*, *eczema* and *lichen planus*. After oral administration gastrointestinal upset and central nervous system toxicities, such as vertigo and excitement, also occur. Consequently, the drug should be used orally only under medical supervision. It is additive with other photosensitizing drugs and the furocumarin pigments in carrots, celery, figs, limes, mustard, parsley and parsnips. It inhibits the metabolism of caffeine.

Dose—*Topical*, as a 1% lotion (see the package literature for details of application and use). *Oral, adults* and *children over 12 yr*, for *vitiligo*, **30** to **40 mg** once a day 2 to 4 hr before exposure to ultraviolet light or at longer than 48-hr intervals 2 or 3 times a week; for *psoriasis*, *mycosis fungoides* or *cutaneous T-cell lymphoma*, **0.6 mg/kg** 2 or 3 hr before UVA exposure (see the package literature for details).

Dosage Forms—Capsules: 10 mg; Lotion: 1%.

Monobenzone

Phenol, 4-(phenylmethoxy)-, Monobenzyl Ether of Hydroquinone; Benoquin (*Elder*)



p-(Benzyloxy)phenol [103-16-2] $C_{13}H_{12}O_2$ (200.24).

Preparation—Prepared in various ways. One method involves condensing sodium *p*-nitrophenolate with benzyl chloride to produce benzyl *p*-nitrophenyl ether followed by (1) reduction of nitro to amino, (2) diazotization of amino and (3) hydrolytic decomposition of the diazonium compound to the corresponding phenol.

Description—White, odorless, crystalline powder possessing very little taste; melts between 117° and 120°.

Solubility—1 g in >10,000 mL water, 14.5 mL alcohol, 29 mL chloroform or 14 mL ether.

Uses—A *depigmenting agent* or *demelanizer*. It acts by interfering with the formation of melanin, which is the principal cutaneous pigment. It is recommended only for the final depigmentation in *vitiligo*. It is not recommended for treatment of lentigo, severe freckling and other types of hyperpigmentation. It is not effective against pigmented moles or malignant melanoma. Its pigment-decreasing action is somewhat erratic. Irritation of varying degrees occurs in a considerable number of patients.

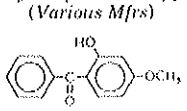
Dose—*Topical, adults* and *children over 12 yr*, to the skin, as a **20%** cream 2 or 3 times a day.

Dosage Forms—Cream: 20%.

Minoxidil—page 837.

Oxybenzone

Methanone, (2-hydroxy-4-methoxyphenyl)phenyl-, (*Various Mfrs*)



2-Hydroxy-4-methoxybenzophenone [131-57-7] $C_{14}H_{12}O_3$ (228.25).

Preparation—Benzoic acid is condensed with resorcinol monomethyl ether by heating in the presence of $ZnCl_2$ or polyphosphoric acid (103% H_3PO_4 equivalent), and PCl_3. US Pat 3,073,866.

Description—White to off-white powder; congeals not lower than 62°.

Solubility—Practically insoluble in water; freely soluble in alcohol or toluene.

Uses—A *sunscreen* with a high molar absorptivity (20,381 at 290 nm), and it absorbs in both the long and short UV spectrum 270–350 nm. Therefore, it serves not only to prevent sunburn but also to protect against the photodynamic, photosensitizing and phototoxic effects of various drugs. Contact with the eyes should be avoided. At present, it is marketed only in combination with other sunscreens.

Dose—*Topical*, as a 3 to 5% cream, 0.5% lipstick and 2 or 3% lotion in combination with other sunscreens.

Ringer's Irrigation—RPS-16, page 762.

Sodium Bicarbonate—page 777.

Sodium Fluoride

Sodium fluoride [7681-49-4] NaF (41.99).

Preparation—By interaction of 40% HF with an equivalent quantity of NaOH or Na_2CO_3.

Description—White, odorless powder.

Solubility—1 g in 25 mL water; insoluble in alcohol.

Uses—A *dental caries prophylactic*. Fluoridation of municipal water supplies is considered a safe and practical public health measure, a concentration of about 1 ppm of fluoride in the water supply resulting in a 50 to 65% reduction in the incidence of dental caries in permanent teeth. Ingested fluoride is effective only while teeth are being formed. The fluoride is incorporated into tooth salts as fluoroapatite. Excessive intake during development of teeth may cause mottling; hence, mottling of newly erupted teeth is an indication to reduce fluoride intake. Where drinking water contains less than 0.7 ppm of fluoride, dietary supplements for children with unerupted teeth may provide some future protection.

Topical application results in changes only in the outer layers of enamel or exposed dentin. In children, repeated application of a 2% solution of the drug to cleaned teeth results in a 16 to 49% reduction of dental caries; adult teeth are protected to a lesser extent by topical application. Topical application also is used to *densensitize* teeth.

Orally administered, it produces new bone formation in some patients with osteoporosis, especially when calcium and vitamin D (and estrogens in women) are administered concomitantly to facilitate mineralization of the new bone. However, the bone may become brittle.

It removes calcium from tissues and also poisons certain enzymes. Large oral doses may cause nausea and vomiting, which usually can be prevented by taking the substance with food. Pastes, rinses, solutions and gels for topical applications should not be swallowed.

Dose (as sodium fluoride)—*Topical*, to the teeth, as a 0.02 to 2% solution, 1.1 or 2.71% gel or 0.22 to 2.3% toothpaste. *Oral*, 1.5 to 3 ppm (equivalent to 0.7 to 1.3 ppm of fluoride ion) in drinking water; as a supplement, when the drinking water contains less than 0.3 ppm of fluoride ion, 0.55 mg a day for infants from 2 wk to 2 yr of age, 1.1 mg once a day for children from 2 to 3 yr and 2.2 mg for those from 3 to 13 yr, and when the drinking water contains 0.3 to 0.7 ppm of fluoride ion, 550 μg once a day for children 2 to 3 yr and 1.1 mg for those 3 to 13 yr. The fluoride ion equivalents of 550 μg, 1.1 mg, and 2.2 mg of the drug are 250 μg, 500 μg, and 1 mg, respectively. For *osteoporosis*, up to 60 mg a day. *Caution: It is poisonous.*

Dosage Forms—Drops: 0.275, 0.55 and 1.1 mg/drop; Gel: 1.1 and 2.71%; Rinse: 0.02, 0.05, 0.2 and 0.44%; Solution: 1.1, 3.3, 5.5, and 20 mg/mL; Chewable Tablets: 0.55, 1.1 and 2.2 mg. Sodium Fluoride and Orthophosphoric Acid: Gel: 1.23% fluoride ion and 1% phosphoric acid.

Sodium Monofluorophosphate

Phosphorofluoridic acid, sodium salt

$FPO(ONa)_2$

Disodium phosphorofluoridate [10163-15-2] (143.95).

Preparation—Substantially pure drug is produced by fusing a mixture of sodium metaphosphate and sodium fluoride, in stoichiometric proportion, in a closed vessel from which moist air is excluded.

Description—White to slightly gray, odorless powder.

Solubility—Freely soluble in water.

Uses—Like *Sodium Fluoride*, above, it promotes the replacement of the hydroxyapatite by fluoroapatite in the tooth salts and, hence, is used as a *dental prophylactic* against dental caries. It has the advantage over sodium fluoride in that the teeth do not require special preparation before application, it is effective when included in dentifrices and in dentifrices there is no hazard with respect to local toxicity to the gingivae or systemic intoxication from ingestion.

Dose—*Topical*, to the teeth, in dentifrice containing 0.76%.

Stannous Fluoride

Tin Difluoride; Fluoristan

Tin fluoride (SnF_2) [7783-47-3] (156.69); contains not less than 71.2% Sn^{2+} (stannous tin), and about 24% F^- (fluoride).

Preparation—Stannous oxide is dissolved in 40% HF and the solution is evaporated out of contact with air.

Description—White, crystalline powder with a bitter, salty taste; melts at about 213°.

Solubility—Freely soluble in water; practically insoluble in alcohol, ether or chloroform.

Uses—Alters the composition and crystalline structure of the hydroxyapatite-like salts that make up the bulk of enamel and dentin, so that the tooth material is more resistant to acidic erosion and dental caries (decay). The substance is applied only topically, so that the tooth substance is only affected in the superficial layers, and it must be applied periodically. It is most effective when applied to the tooth surface after the teeth have been cleaned thoroughly by a dentist. However, there is good evidence that even when incorporated into tooth pastes the drug has a retardant effect on the development of dental caries.

Dose—*Topical*, to the teeth, generally as 0.4% gel or 0.1% rinse.

Dosage Forms—Capsules (for solution): 0.4, 0.65 and 0.8 g; Concentrate: 30%; Gel: 0.4%.

Titanium Dioxide

Titanic Anhydride

Titanium oxide (TiO_2) [13463-67-7] TiO_2 (79.88).

Preparation—By adding ammonia or an alkali carbonate to a solution of titanyl sulfate ($TiOSO_4$). Titanic acid [$Ti(OH)_4$ or $TiO(OH)_2$] is precipitated and, after filtration and washing, is dried and ignited.

Description—White, amorphous, tasteless, odorless, infusible powder; density about 4; suspension in water (1 in 10) neutral to litmus.

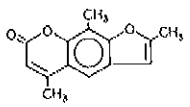
Solubility—Insoluble in water, HCl, HNO_3 or dilute H_2SO_4.

Uses—Its powder has a very high reflectance at visible and UV wavelengths, and, hence, it serves as an excellent white pigment. In ointments or lotions it reflects a very high proportion of incident sunlight, hence, protecting the skin from sunburn and serving as a *sunblock*. It also is used in cosmetics and as a dusting powder. Topically, it is devoid of toxicity.

Dose—*Topical*, as 2 to 25% cream, lotion or ointment as required.

Trioxsalen

7*H*-Furo[3,2-*g*][1]benzopyran-7-one, 2,5,9-trimethyl-, 6-Hydroxy-β,2,7-trimethyl-5-benzofuranacrylic Acid δ-Lactone; Trisoralen (*Elder*)



[3902-71-4] $C_{14}H_{12}O_3$ (228.25).

Caution: Avoid contact with the skin.

Preparation—2-Methylresorcinol is cyclized with ethyl acetoacetate with the aid of sulfuric acid to 7-hydroxy-4,8-dimethylcoumarin (I). Treatment with allyl bromide in the presence of potassium carbonate transforms I into the 7-allyloxy compound which, on reacting with acetic anhydride in the presence of *N,N*-diethylaniline and anhydrous sodium acetate, rearranges and esterifies to give the 7-acetoxy-6-allyl compound (II). Bromination of II followed by reaction with sodium methoxide yields trioxsalen. US Pat 3,201,421.

Description—White to off-white, odorless, tasteless crystalline solid; stable in light, air and heat; melts at about 230°.

Solubility—1 g in 1150 mL alcohol, 84 mL chloroform or 43 mL methylenedichloride; practically insoluble in water.

Uses—Although not a topical drug, it closely relates to other drugs in this section. It facilitates the action of near UV light to induce melanin (skin pigment) formation. It is used to cause repigmentation in idiopathic *vitiligo* and to enhance pigmentation to *increase tolerance to sunlight* or for *cosmetic purposes*. The increased tolerance to sunlight does not occur until enhanced pigmentation has occurred, and the user must be cautioned that severe sunburning with less than normal exposure can occur early during the course of treatment. The increase in dermal pigment occurs gradually over a period of several days of repeated exposure. Care must be taken to protect the eyes and lips during treatment. The manufacturer's recommended schedule of exposure should be used except at high altitudes, where exposure times should be appropriately reduced.

It is contraindicated in persons with photosensitizing diseases, such as infectious leukoderma, porphyria or lupus erythematosus and when photosensitizing drugs are being given. The drug sometimes may cause gastric irritation and emesis. Children under 12 should not take it.

Dose—*Oral, adults* and *children over 12 yr*, **5 to 10 mg** 2 hr before exposure to sunlight. For the treatment of vitiligo the exposure should be repeated once a day for 4 days, and subsequent exposures should be determined according to the results of the initial 4 days. For the enhancement of pigmentation, treatment should not exceed 2 weeks, and the total accumulated dose in any one treatment course should not exceed 140 mg. Persons who show side effects of the drug should take only 5 mg; the duration of use will be necessarily prolonged over that in persons taking the usual dose of 10 mg.

Dosage Forms—Tablets: 5 mg.

Urea—page 931.

Other Miscellaneous Topical Drugs

Allantoin 2,5-Dioxo-4-imidazolidinylurea [97-59-6]; $C_4H_6N_4O_3$ (158.12)]—Prepared by oxidation of uric acid. Colorless crystals melting at 238°. 1 g dissolves in 190 mL water or 500 mL alcohol; nearly insoluble in ether. *Uses:* In World War I it was noticed that maggot-infested wounds seemed to heal better than uninfested wounds, an effect attributed to this drug produced by maggots. It is used topically as a vulnerary to stimulate tissue repair in suppurating wounds, resistant ulcers, acne, seborrhea, cold sores, hemorrhoids and various dermatologic infections and psoriasis. It frequently is combined with astringents, keratolytics, coal tar, antiseptics and antifungal drugs. The silver salt has been used in the topical treatment of extensive burns. *Dose: Topical*, 0.2 to 2% in creams, lotions or shampoos and 0.3 to 0.5% in ointments for hemorrhoids.

Cinoxate [2-Ethoxyethyl *p*-methoxycinnamate [104-28-9]; $C_{14}H_{18}O_4$ (250.29)]—A viscous liquid that may have a slightly yellow tinge; boils at about 185°. Practically insoluble in water; miscible with alcohols. *Uses:* A *sunscreen* that absorbs UV light at 270 to 328 nm and has a relatively high molar absorptivity (19,400 at 306 nm) but not absorbing well throughout the entire offending range of UV light. Consequently, it is used principally in preparations intended to promote tanning rather than to protect against photosensitivity and phototoxicity. *Dose: Topical*, 1.75 to 4% in creams, gels or lotions.

Dextranomer [Dextran 2,3-dihydroxypropyl-2-hydroxy-1,3-propanediyl ether [56087-11-7] Dextran polymer; Debrisan (*Pharmacia*)]—Small, dry beads of a three-dimensional dextran polymer; highly hygroscopic. 1 g absorbs about 4 g water. *Uses:* For drying, cleansing and debridement of exudative *venous stasis ulcers, infected wounds* and *burns;* it is not useful for cleansing nonexudative wounds or lesions. The beads not only absorb water but also proteins, including fibrin/fibrinogen degradation products and, thus, prevent encrustation. The beads are poured into the cleansed wound, which is circumscribed with petroleum jelly, and a compress is taped in place to retain the material. Changes may be made up to 3 or 4 times a day, as needed. The beads must be removed before skin grafting is attempted. Care must be taken to prevent cross-contamination from patient to patient. On the floor the beads are slippery and, thus, hazardous.

Digalloyl Trioleate [[17048-39-4; 27436-80-2] $C_{68}H_{108}O_{12}$ (1115.59)]—*Uses:* A sunscreen with an absorption band at 270 to 320 nm. It is used topically as a 3.5% cream or 2.5% lipstick.

Dihydroxyacetone [1,3-Dihydroxydimethyl ketone [96-26-4] $C_3H_6O_3$ (90.08)]—The ketone resulting from oxidation of the secondary alcohol group of glycerin. A crystalline powder; fairly hygroscopic; characteristic odor and sweet taste. The normal form is the dimer, slowly soluble in 1 part water or 15 parts alcohol; the monomer formed in solution is very soluble in water, alcohol or ether. *Uses:* Interacts with keratin in the stratum corneum to form a dark pigment that simulates the appearance of a suntan. It is incorporated in several sunscreen preparations. Since the sunscreen component is usually present in a concentration lower than optimal, such preparations may not provide protection to photosensitive persons.

Ethyl Dihydroxypropylaminobenzoate [Ethyl 4-[bis(hydroxypropyl)]aminobenzoate [58882-17-0] $C_{15}H_{23}NO_4$ (281.35); Amerscreen (*Amerchol*)]—*Uses:* A sunscreen with a limited absorption spectrum (280 to 330 nm) characteristic of *p*-aminobenzoates but a relatively high molar absorptivity. It is used mainly in suntan products. *Dose: Topical*, in concentrations of 1 to 5%.

Ethylhexyl Methoxycinnamate [2-Ethylhexyl *p*-methoxycinnamate [5466-77-3] $C_{18}H_{26}O_3$ (290.40)]—*Uses:* A sunscreen with a narrow absorption band of 290 to 320 nm and a moderate molar absorptivity. *Dose: Topical*, in 2 to 7.5% concentration in creams, lotions and oils.

Glyceryl *p*-Aminobenzoate [1,2,3-Propanetriol 1-(4-aminobenzoate) [136-44-7] $C_{10}H_{13}NO_4$ (211.21)]—Prepared by esterification of aminobenzoic acid with glycerin. A waxy semisolid or syrup. Insoluble in water, oils or fats; soluble in ethanol, isopropanol or propylene glycol. *Uses:* A sunscreen that absorbs UV light at 264 to 315 nm and which has a relatively high molar absorptivity (17,197 at 295 nm) but a limited spectrum, therefore used primarily to promote tanning rather than to protect sensitive persons. *Dose: Topical*, 2 to 3% in lotions.

Homosalate [3,3,5-Trimethylcyclohexyl salicylate; homomenthyl salicylate [118-56-9] $C_{16}H_{22}O_3$ (262.36); ing of Coppertone (*Plough*); Filtrosol "A" (*Norda*); Heliophan (*Greeff*)]—*Uses:* A liquid with relatively low molar absorptivity (6,720 at 310 nm) and limited absorption in the near ultraviolet range (290 to 315 nm), so that it is used mainly to promote tanning. Photosensitive persons may not be protected from burns and phototoxicity. *Dose: Topical*, 4 to 10% in creams, lotions or oils.

Methyl Anthranilate [Methyl 2-aminobenzoate [134-20-3] $C_8H_9NO_2$ (151.16)]—A constituent of several essential oils; also obtained by esterifying anthranilic acid with methyl alcohol. A crystalline substance; melts at 25°. Slightly soluble in water; freely soluble in alcohol or ether. *Uses:* A sunscreen, with the lowest molar absorptivity of all sunscreens (941 at 315 nm); also, it does not absorb throughout the near UV range (absorption band, 290 to 320 nm) and, therefore, is used in combination with other sunscreens or light-protectives. It also is used as a perfume in ointments and cosmetics. *Dose: Topical*, to the skin, 5% in creams, lotions or ointment(s.

Octyl salicylate—*Uses:* A sunscreen with an absorption band at 280 to 320 nm and a moderate absorptivity. It is used primarily in conjunction with other sunscreens in suntan products.

Padimate A [Pentyl *p*-(dimethylamino)benzoate [14779-78-3] $C_{14}H_{21}NO_2$ (235.33); (*Various Mfrs*)]—A mixture of pentyl, isopentyl and 2-methylbutyl esters of *p*-aminobenzoic acid. Yellow liquid with a faint, aromatic odor. Practically insoluble in water or glycerin; soluble in alcohol, chloroform, isopropyl alcohol or mineral oil. *Uses:* A sunscreen of moderate molar absorptivity but relatively narrow UV absorption spectrum (290 to 315 nm) characteristic of other aminobenzoic acid derivatives. *Dose: Topical*, to the skin, as a 1.4 to 8% cream, foam, lotion or stick.

Padimate O [2-Ethylhexyl 4-(dimethylamino)benzoate [21245-02-3] $C_{17}H_{27}NO_2$ (277.41); (*Various Mfrs*)]—A light-yellow mobile liquid with a faint, aromatic odor. Practically insoluble in water, alcohol or mineral oil. *Uses:* See *Padimate A*.

Red Petrolatum—*Uses:* Owing to its opacity, it is used in sunblock creams, ointments and sticks. Concentrations range from 30 to 100%.

CHAPTER 66

Pharmaceutical Necessities

Ewart A Swinyard, PhD, DSc (Hon)
Professor Emeritus of Pharmacology
College of Pharmacy and School of Medicine
University of Utah
Salt Lake City, UT 84112

Werner Lowenthal, PhD
Professor of Pharmacy and Pharmaceutics and Professor of Educational Planning and Development
School of Pharmacy
Virginia Commonwealth University
Richmond, VA 23298

This chapter describes substances that are of little or no therapeutic value, but which are useful in the manufacture and compounding of various pharmaceutical preparations. Hence, they are referred to as pharmaceutical necessities. The substances described include antioxidants and preservatives; coloring, flavoring and diluting agents; emulsifying and suspending agents; ointment bases; pharmaceutical solvents and miscellaneous agents. For a more detailed review of the uses of these agents, the interested reader is referred to the various chapters in Part 8 of this book.

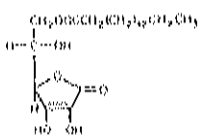
Antioxidants and Preservatives

An antioxidant is a substance capable of inhibiting oxidation and that may be added for this purpose to pharmaceutical products subject to deterioration by oxidative processes as, for example, the development of rancidity in oils and fats or the inactivation of some medicinals in the environment of their dosage forms. A preservative is, in the common pharmaceutical sense, a substance that prevents or inhibits microbial growth and may be added to pharmaceutical preparations for this purpose to avoid consequent spoilage of the preparations by microorganisms. Both antioxidants and preservatives have many applications in making medicinal products.

Alcohol—page 1314.

Ascorbyl Palmitate

L-Ascorbic acid, 6-hexadecanoate; Ascorbic Acid Palmitate (ester)



L-Ascorbic acid 6-palmitate [137-66-6] $C_{22}H_{38}O_7$ (414.54).

Preparation—By condensing palmitoyl chloride with ascorbic acid in the presence of a suitable dehydrochlorinating agent such as pyridine.

Description—White to yellowish white powder having a characteristic odor; melts 107° and 117°.

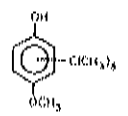
Solubility—1 g in >1000 mL of water, 125 mL of alcohol, >1000 mL of chloroform or >3000 mL of ether.

Uses—An *antioxidant* used in foods and pharmaceuticals. It also is used to prevent rancidity, to prevent the browning of cut apples, in meat curing and in the preservation of canned or frozen foods.

Benzoic Acid—page 1235.
Benzalkonium Chloride—page 1104.
Benzethonium Chloride—page 1170.
Benzyl Alcohol—page 1056.

Butylated Hydroxyanisole

Phenol, (1,1-dimethylethyl)-4-methoxy-; Tenox BHA (*Eastman*)



tert-Butyl-4-methoxyphenol [25013-16-5] $C_{11}H_{16}O_2$ (180.25).

Preparation—By an addition interaction of *p*-methoxyphenol and 2-methylpropene. US Pat 2,428,745.

Description—White or slightly yellow, waxy solid having a faint, characteristic odor.

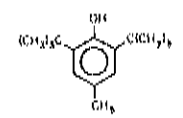
Solubility—Insoluble in water; 1 g in 4 mL of alcohol, 2 mL of chloroform or 1.2 mL of ether.

Uses—An *antioxidant* in cosmetics and pharmaceuticals containing fats and oils.

Butylparaben—page 1170.

Butylated Hydroxytoluene

Phenol, 2,6-bis(1,1-dimethylethyl)-4-methyl-, Butylated Hydroxytoluene Crystalline (*Diamond-Shamrock*); Tenox BHT (*Eastman*)



2,6-Di-*tert*-butyl-*p*-cresol [128-37-0] $C_{15}H_{24}O$ (220.35).

Preparation—By an addition interaction of *p*-cresol and 2-methylpropene. US Pat 2,428,745.

Description—White, tasteless crystals with a mild odor; stable in light and air; melts at 70°.

Solubility—Insoluble in water; 1 g in 4 mL of alcohol, 1.1 mL of chloroform or 1.1 mL of ether.

Uses—An *antioxidant* employed to retard oxidative degradation of oils and fats in various cosmetics and pharmaceuticals.

Cetylpyridinium Chloride—page 1171.

Chlorobutanol

2-Propanol, 1,1,1-trichloro-2-methyl-, Chlorbutol; Chlorbutanol; Acetone chloroform; Chloretone (*Parke-Davis*)

$(CCl_3)C(CH_3)_2OH$

1,1,1-Trichloro-2-methyl-2-propanol [57-15-8] $C_4H_7Cl_3O$ (177.46); *hemihydrate* [6001-64-5] (186.46).

Preparation—Chloroform undergoes chemical addition to acetone under the catalytic influence of powdered potassium hydroxide.

Description—Colorless to white crystals, of a characteristic, somewhat camphoraceous odor and taste; anhydrous melts about 95°; hydrous melts about 76°; boils with some decomposition 165 and 168°.

Solubility—1 g in 125 mL of water, 1 mL of alcohol or about 10 mL of glycerin; freely soluble in chloroform, ether or volatile oils.

Incompatibilities—The anhydrous form must be used in order to prepare a clear solution in liquid petrolatum. It is decomposed by *alkalies; ephedrine* is sufficiently alkaline to cause its breakdown with the formation of ephedrine hydrochloride which will separate from a liquid petrolatum solution. It is only slightly soluble in water, hence alcohol must be used to dissolve the required amount in certain vehicles. A soft mass is produced by trituration with *antipyrine, menthol, phenol* and other substances.

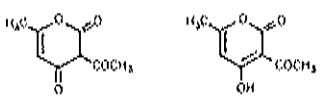
Uses—Topically, as a solution in clove oil as a *dental analgesic*. It has *local anesthetic* potency to a mild degree and has been employed as an anesthetic dusting powder (1 to 5%) or ointment (10%). It has antibacterial and germicidal properties. It is used chiefly as a *preservative* in solutions of epinephrine, posterior pituitary, etc. When administered orally, it has much the same therapeutic use as chloral hydrate. Hence, it has been employed as a sedative and hypnotic. It has been taken orally to allay vomiting due to gastritis.

Dose—*Topical*, as a 25% solution in clove oil.

Other Dose Information—The oral dose is 300 mg to 1 g, given in tablets or capsules.

Dehydroacetic Acid

Keto form: 2*H*-Pyran-2,4(3*H*)-dione, 3-acetyl-6-methyl-,



Enol form: 3-Acetyl-4-hydroxy-6-methyl-2H-pyran-2-one [520-45-6 (Keto)], [771-03-9 (enol)] $C_8H_8O_4$ (168.15).

Preparation—By fractional distillation of a mixture of ethyl acetoacetate and sodium bicarbonate, maintaining almost total reflux conditions, allowing only ethanol to be removed. The residue is distilled under vacuum. *Org Syn Coll Vol III:* 231, 1956.

Description—White to creamy-white crystalline powder melting about 110° with sublimation.

Solubility—One g dissolves in 25 g of acetone, 18 g of benzene, 5 g of methanol or 3 g of ethanol.

Uses—Preservative.

Ethylenediamine

1,2-Ethanediamine

$H_2NCH_2CH_2NH_2$

Ethylenediamine [107-15-3] $C_2H_8N_2$ (60.10).

Caution—Use care in handling because of its caustic nature and the irritating properties of its vapor.

Note—It is strongly alkaline and may readily absorb carbon dioxide from the air to form a nonvolatile carbonate. Protect it against undue exposure to the atmosphere.

Preparation—By reacting ethylene dichloride with ammonia, then adding NaOH and distilling.

Description—Clear, colorless or only slightly yellow liquid, having an ammonia-like odor and strong alkaline reaction; miscible with water and alcohol; anhydrous boils 116 to 117° and solidifies at about 8°; volatile with steam; a strong base and readily combines with acids to form salts with the evolution of much heat.

Uses—A *pharmaceutical necessity* for *Aminophylline Injection*. It is irritating to skin and mucous membranes. It also may cause sensitization characterized by asthma and allergic dermatitis.

Ethylparaben—page 1171.
Ethyl Vanillin—page 1204.
Glycerin—page 1027.
Hypophosphorus Acid—page 1322.
Methylparaben—page 1172.

Monothioglycerol

1,2-Propanediol, 3-mercapto-,

$HSCH_2CH(OH)CH_2OH$

3-Mercapto-1,2-propanediol [96-27-5] $C_3H_8O_2S$ (108.15).

Preparation—An ethanolic solution of 3-chloro-1,2-propanediol is heated with potassium bisulfide.

Description—Colorless or pale yellow, viscous liquid having a slight sulfidic odor; hygroscopic; specific gravity 1.241 to 1.250; pH (1 in 10 solution) 3.5 to 7.

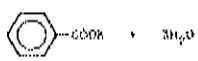
Solubility—Freely soluble in water; miscible with alcohol; insoluble in ether.

Uses—A pharmaceutic aid stated to be used as a preservative. It has been used in 1:5000 solution to stimulate healing of wounds, and as a 1:1000 jelly in atrophic rhinitis.

Phenol—page 1323.
Phenylethyl Alcohol—page 1297.
Phenylmercuric Nitrate—page 1172.

Potassium Benzoate

Benzoic acid, potassium salt



[582-25-2] $C_7H_5KO_2$ (160.21) (anhydrous).

Description—Crystalline powder.

Solubility—Soluble in water or alcohol.

Uses—Preservative.

Potassium Metabisulfite

Dipotassium pyrosulfite

[16731-55-8] $K_2S_2O_5$ (222.31).

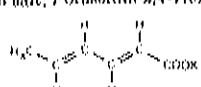
Description—White crystals or crystalline powder with an odor of SO_2. Oxidizes in air to the sulfate. May ignite on powdering in a mortar if too much heat develops.

Solubility—Freely soluble in water; insoluble in alcohol.

Uses—Antioxidant.

Potassium Sorbate

2,4-Hexadienoic acid, (*E,E*)-, potassium salt; 2,4-Hexadienoic acid, potassium salt; Potassium 2,4-Hexadienoate



Potassium (*E,E*)-sorbate; potassium sorbate [590-00-1] [24634-61-5] $C_6H_7KO_2$ (150.22).

Preparation—Sorbic Acid is reacted with an equimolar portion of KOH. The resulting potassium sorbate may be crystallized from aqueous ethanol. US Pat 3,173,948.

Description—White crystals or powder with a characteristic odor; melts about 270° with decomposition.

Solubility—1 g in 4.5 mL of water, 35 mL of alcohol, >1000 mL of chloroform or >1000 mL of ether.

Uses—A water-soluble salt of sorbic acid used in pharmaceuticals to *inhibit the growth of molds and yeasts*. Its toxicity is low, but it may irritate the skin.

Propylparaben—page 1173.

Sassafras Oil—page 1300.

Sodium Benzoate—page 1173.

Sodium Bisulfite

Sulfurous acid, monosodium salt; Sodium Hydrogen Sulfite; Sodium Acid Sulfite; Leucogen

Monosodium sulfite [7631-90-5] $NaHSO_3$ and sodium metabisulfite ($Na_2S_2O_5$) in varying proportions; yields 58.5–67.4% of SO_2.

Description—White or yellowish white crystals or granular powder having the odor of sulfur dioxide; unstable in air.

Solubility—1 g in 4 mL of water; slightly soluble in alcohol.

Uses—An *antioxidant* and *stabilizing agent*. Epinephrine hydrochloride solutions may be stabilized by the addition of small quantities of the salt. It also is used to help solubilize kidney stones. It is useful for removing permanganate stains and for solubilizing certain dyes and other chemicals (see *Monadione Sodium Bisulfite*, RPS-17, page 1011).

Sodium Metabisulfite

Disulfurous acid, disodium salt

Disodium pyrosulfite [7681-57-4] $Na_2S_2O_5$ (190.10).

Preparation—Formed when sodium bisulfite undergoes thermal dehydration. It also may be prepared by passing sulfur dioxide over sodium carbonate.

Description—White crystals or white to yellowish crystalline powder having an odor of sulfur dioxide; on exposure to air and moisture, it is slowly oxidized to sulfate.

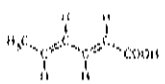
Solubility—1 g in 2 mL of water; slightly soluble in alcohol; freely soluble in glycerin.

Uses—A *reducing agent*. It is used in easily oxidized pharmaceuticals, such as epinephrine hydrochloride and phenylephrine hydrochloride injections, to retard oxidation.

Sodium Propionate—page 1230.

Sorbic Acid

2,4-Hexadienoic acid, (*E,E*)-; 2,4-Hexadienoic acid



(*E,E*)-Sorbic acid; Sorbic acid [22500-92-1] [110-44-1] $C_6H_8O_2$ (112.13).

Preparation—By various processes. Refer to US Pat 2,921,090.

Description Free-flowing, white, crystalline powder, having a characteristic odor; melts about 133°.

Solubility 1 g in 1000 mL of water, 10 mL of alcohol, 16 mL of chloroform, 30 mL of ether or 19 mL of propylene glycol.

Uses—A *mold and yeast inhibitor*. It also is used as a fungistatic agent for foods, especially cheeses.

Sulfur Dioxide

Sulfur dioxide [7446-09-5] SO_2 (64.06).

Preparation—By burning sulfur or sulfides and by reacting a bisulfite or a sulfite with a strong acid.

Description—Colorless, nonflammable gas, with a strong, suffocating, odor characteristic of burning sulfur; 1 L weighs 2.927 g at 760 mm and 0°; readily liquefies under pressure forming a colorless liquid with a density of approximately 1.5 g/mL and a boiling point of −10°.

Solubility—1 volume of water dissolves approximately 36 volumes of it at 760 mm and 20°; 1 volume of alcohol dissolves approximately 114 volumes under the same conditions; soluble in ether or chloroform.

Note—It is used mostly in the form of a gas in pharmaceutical applications, and is described herein for such purposes. However, it is usually packaged under pressure, hence the USP specifications (Water, Nonvolatile residue and *Sulfuric acid), are designed for the testing of its liquid form.*

Uses—The gas in the presence of moisture forms sulfurous acid which is a *bleaching agent*, *fungicide* and *bactericide*. For this reason fruits often are exposed to the gas before drying to prevent darkening and the growth of molds and bacteria. The gas is also an *antioxidant* and a pharmaceutical necessity for *Injections*. It may be intensely irritating to the eyes and respiratory tract.

Thimerosal—page 1173.

Other Antioxidants and Preservatives

Anoxomer [1,4-Benzenediol, 2-(1,1-dimethylethyl)-, polymer with diethenyl benzene, 4-(1,1-dimethylethyl)phenol, 4-methoxyphenol, 4,4′-(1-methylethylidene)bis[phenol] and 4-methylphenol {80837-57-2} $(C_{10}H_{14}O_2)_n(C_{10}H_{10})_x(C_{10}H_{14}O)_n(C_7H_8O_2)_x(C_{16}H_{16}O_2)_y(C_7H_8O)_z$. *Uses:* Antioxidant and food additive.

Maleic Acid BP [*cis*-Butanedioic acid $C_4H_4O_4$ (116.07); Toxilic acid] —*Preparation:* Benzene vapor is oxidized by passage over heated vanadium pentoxide. Odorless, white, crystalline powder having a strongly acid taste; melts about 136°. Soluble in 1.6 parts of water, 2 parts of alcohol or 12 parts of ether. *Uses:* In the preparation of ergometrine maleate injection or as a rancidity retardant in fats and oils (1:10,000).

Propyl Gallate BP [Propyl 3,4,5-Trihydroxybenzoate]—White to creamy-white crystalline powder; odorless; slightly bitter taste. Soluble in 1000 parts of water or 3 parts of alcohol. *Uses:* A preservative.

Coloring, Flavoring and Diluting Agents

The use of properly colored and flavored medicinal substances, although offering no particular therapeutic advantage, is of considerable importance psychologically. A water-clear medicine is not particularly acceptable to most patients, and, in general, is thought to be inert. Many very active medicinal substances are quite unpalatable, and the patient may fail to take the medicine simply because the taste or appearance is objectionable. Disagreeable medication can be made both pleasing to the taste and attractive by careful selection of the appropriate coloring, flavoring and diluting agents. Therefore, judicious use of these substances is important in securing patient cooperation in taking or using the prescribed medication and continued compliance with the prescriber's intent.

Coloring Agents or Colorants

Coloring agents may be defined as compounds employed in pharmacy solely for the purpose of imparting color. They may be classified in various ways, eg, inorganic or organic. For the purpose of this discussion two subdivisions are used: *Natural Coloring Principles* and *Synthetic Coloring Principles*. The members of these groups are used as colors for pharmaceutical preparations, cosmetics, foods and as bacteriological stains and diagnostic agents.

Natural Coloring Principles

Natural coloring principles are obtained from mineral, plant and animal sources. They are used primarily for artistic purposes, as symbolic adornments of natives, as colors for foods, drugs and cosmetics and for other psychological effects.

Mineral colors frequently are termed *pigments* and are

used to color lotions, cosmetics and other preparations, usually for external application. Examples are *Red Ferric Oxide* (page 1328) and *Yellow Ferric Oxide* (page 1328), titanium dioxide (page 772) and carbon black.

The term pigment also is applied generically to plant colors by phytochemists. Many plants contain coloring principles that may be extracted and used as colorants, eg, chlorophyll. Anattenes are obtained from annatto seeds and give yellow to orange water-soluble dyes. Natural beta-carotene is a yellow color extracted from carrots and used to color margarine. Alizarin is a reddish-yellow dye obtained from the madder plant. The indigo plant is the source of a blue pigment called indigo. Flavones, such as riboflavin, rutin, hesperidin and quercetin, are yellow pigments. Saffron is a glycoside that gives a yellow color to drugs and foods. Cudbear and red saunders are two other dyes obtained from plants. Most plant colors now have been characterized and synthesized, however, and those with the desirable qualities of stability, fastness and pleasing hue are available commercially as synthetic products.

Animals have been a source of coloring principles from the earliest periods of recorded history. For example, *Tyrian purple*, once a sign of royalty, was prepared by air oxidation of a colorless secretion obtained from the glands of a snail (*Murex brandaris*). This dye now is known to be 6,6'-dibromoindigo, and has been synthesized, but cheaper dyes of the same color are available. Cochineal from the insect *Coccus cacti* contains the bright-red coloring principle *carminic acid*, a derivative of anthraquinone. This dye is no longer used in foods and pharmaceuticals due to *Salmonella* contamination.

Synthetic Coloring Principles

Synthetic coloring principles date from 1856 when W H Perkin accidentally discovered *mauveine*, also known as a *Perkin's purple*, while engaged in unsuccessful attempts to synthesize quinine. He obtained the dye by oxidizing aniline containing *o*- and *p*-toluidines as impurities. Other discoveries of this kind followed soon after, and a major industry grew up in the field of coal-tar chemistry.

The earliest colors were prepared from aniline and for many years all coal-tar dyes were called aniline colors, irrespective of their origin. The coal-tar dyes include more than a dozen well-defined groups among which are *nitroso-dyes, nitro-dyes, azo-dyes, oxazines, thiazines, pyrazolones, xanthenes, indigoids, anthraquinones, aeridines, rosanilines, phthaleins, quinolines* and others. These in turn are classified, according to their method of use, as *acid dyes* and *basic dyes*, or *direct dyes* and *mordant dyes*.

Certain structural elements in organic molecules, called chromophore groups, give color to the molecules, eg, azo (—N=N—), nitroso (—N=O), nitro (—NO_2), azoxy (—N=N—O—), carbonyl (>C=O) and ethylene (>C=C<). Other such elements augment the chromophore groups, eg, methoxy, hydroxy and amino groups.

Stability—Most dyes are relatively unstable chemicals due to their unsaturated structures. They are subject to fading due to light, metals, heat, microorganisms, oxidizing and reducing agents plus strong acids and bases. In tablets, fading may appear as spotting and specking.

Uses—Most synthetic coloring principles are used in coloring fabrics and for various artistic purposes. They also find application as indicators, bacteriological stains, diagnostic aids, reagents in microscopy, etc.

Many coal-tar dyes originally were used in foodstuffs and beverages without careful selection or discrimination between those that were harmless and those that were toxic and without any supervision as to purity or freedom from poisonous constituents derived from their manufacture.

After the passage of the Food and Drugs Act in 1906, the US Department of Agriculture established regulations by which a few colors came to be known as *permitted colors*. Certain of these colors may be used in foods, drugs and cosmetics, but only after certification by the FDA that they meet certain specifications. From this list of permitted colors may be produced, by skillful blending and mixing, other colors that may be used in foods, beverages and pharmaceutical preparations. Blends of certified dyes must be recertified.

The word "permitted" is used in a restricted sense. It does not carry with it the right to use colors for purposes of deception, even though they are "permitted" colors, for all food laws have clauses prohibiting the coloring of foods and beverages in a manner so as to conceal inferiority or to give a false appearance of value.

The certified colors are classified into three groups: FD&C dyes which legally may be used in foods, drugs and cosmetics, D&C dyes which legally may be used in drugs and cosmetics and External D&C dyes which legally may be used only in externally applied drugs and cosmetics. There are specific limits for the pure dye, sulfated ash, other extractives, soluble and insoluble matter, uncombined intermediates, oxides, chlorides and sulfates. As the use status of these colors is subject to change, the latest regulations of the FDA should be consulted to determine how they may be used—especially since several FD&C dyes formerly widely used have been found to be carcinogenic even when "pure" and, therefore, have been banned from use.

The Coal-Tar Color Regulations specify that the term "externally applied drugs and cosmetics" means drugs and cosmetics which are applied only to external parts of the body and not to the lips or any body surface covered by mucous membrane. No certified dye, regardless of its category, legally may be used in any article which is to be applied to the area of the eye.

Lakes are calcium or aluminum salts of certified dyes extended on a substrate of alumina. They are insoluble in water and organic solvents, hence are used to color powders, pharmaceuticals, foods, hard candies and food packaging.

The application of dyes to pharmaceutical preparations is an art that can be acquired only after an understanding of the characteristics of dyes and knowledge of the composition of the products to be colored has been obtained. Specific rules for the choice or application of dyes to pharmaceutical preparations are difficult to formulate. Each preparation may present unique problems.

Preparations which may be colored include most liquid pharmaceuticals, powders, ointments and emulsions. Some general hints may be offered in connection with solutions and powders, but desired results usually can be obtained only by a series of trials. In general, an inexperienced operator tends to use a much higher concentration of the dye than is necessary, resulting in a dull color. The amount of dye present in any pharmaceutical preparation should be of a concentration high enough to give the desired color and low enough to prevent toxic reactions and permanent staining of fabrics and tissues.

Liquids (Solutions)—The dye concentration in liquid preparations and solutions usually should come within a range of 0.0005% (1 in 200,000) and 0.001% (1 in 100,000), depending upon the depth of color wanted and the thickness of column to be viewed in the container. With some dyes, concentrations as low as 0.0001% (1 in 1,000,000) may have a distinct tinting effect. Dyes are used most conveniently in the form of stock solutions.

Powders—White powders usually require the incorporation of 0.1% (1 in 1000) of a dye to impart a pastel color. The dyes may be incorporated into the powder by dry-blending in a ball mill or, on a small scale, with a mortar and pestle.

The dye is incorporated by trituration and geometric dilution. Powders also may be colored evenly by adding a solution of the dye in alcohol or some other volatile solvent having only a slight solvent action on the powder being colored. When this procedure is employed, the solution is added in portions, with thorough mixing after each addition, after which the solvent is allowed to evaporate from the mixture.

Many of the syrups and elixirs used as flavoring and diluting agents are colored. When such agents are used no further coloring matter is necessary. The use of colored flavoring agents is discussed in a subsequent section. However, when it is desired to add color to an otherwise colorless mixture, one of the agents described in the first section may be used.

Incompatibilities—FD&C dyes are mainly anionic (sodium salts), hence are incompatible with cationic substances. Since the concentrations of these substances are generally very low, no precipitate is evident. Polyvalent ions such as calcium, magnesium and aluminum also may form insoluble compounds with dyes. A pH change may cause the color to change. Acids may release the insoluble acid form of the dye.

Caramel

Burnt Sugar Coloring

A concentrated solution of the product obtained by heating sugar or glucose until the sweet taste is destroyed and a uniform dark brown mass results, a small amount of alkali, alkaline carbonate or a trace of mineral acid being added while heating.

Description—Thick, dark brown liquid with the characteristic odor of burnt sugar, and a pleasant, bitter taste; specific gravity not less than 1.30; 1 part dissolved in 1000 parts of water yields a clear solution having a distinct yellowish orange color which is not changed and no precipitate is formed after exposure to sunlight for 6 hr; when spread in a thin layer on a glass plate, it appears homogeneous, reddish brown and transparent.

Solubility—Miscible with water in all proportions and with dilute alcohol up to 55% by volume; immiscible with ether, chloroform, acetone, benzene, solvent hexane or turpentine oil.

Uses—To produce a brown color in elixirs, syrups and other preparations.

Flavoring Agents

Flavor

The word flavor refers to a mixed sensation of taste, touch, smell, sight and sound, all of which combine to produce an infinite number of gradations in the perception of a substance. The four primary tastes—*sweet*, *bitter*, *sour* and *saline*—appear to be the result partly of physicochemical and partly of psychological action. Taste buds (Fig 66-1), located mainly on the tongue, contain very sensitive nerve endings that react, in the presence of moisture, with the flavors in the mouth and as a result of physicochemical activity electrical impulses are produced and transmitted via the seventh, ninth and tenth cranial nerves to the areas of the brain which are devoted to the perception of taste. Some of the taste buds are specialized in their function, giving rise to areas on the tongue which are sensitive to only one type of taste. The brain, however, usually perceives taste as a composite sensation, and accordingly the components of any flavor are not readily discernible. Children have more taste buds than adults, hence are more sensitive to tastes.

Taste partly depends on the ions which are produced in the mouth, but psychologists have demonstrated that sight (color) and sound also play a definite role when certain reflexes become conditioned through custom and association of sense perceptions. Thus, in the classic experiments of Pavlov demonstrating "conditioned reflexes," the ringing of a bell or the showing of a circle of light caused the gastric juices of a dog to flow although no food was placed before it, and much of the enjoyment derived from eating celery is due to its crunchy crispness as the fibrovascular bundles are crushed. The effect of color is just as pronounced; oleomargarine is unpalatable to most people when it is uncolored, but once the dye has been incorporated gourmets frequently cannot distinguish it from butter. Color and taste must coincide, eg, cherry flavor is associated with a red color.

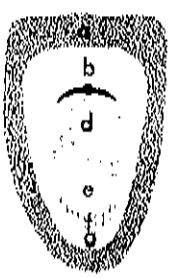


Fig 66-1. Upper Surface of the tongue. a: Taste receptors for all tastes; b: sweet, salty and sour; c: salty and sour; d: sour only; e: no taste sensation; f: sweet and sour and g: bitter, sweet and sour (adapted from Crocker EC: *Flavor*, McGraw-Hill, New York, 22, 1945).

A person suffering from a head cold finds his food much less palatable than usual because his sense of smell is impaired, and, if the nostrils are held closed, raw onions taste sweet and it is much easier to ingest castor oil and other nauseating medicines. The volatility of a substance is an important factor that is influenced by the warmth and moisture of the mouth since the more volatile a compound, the more pronounced its odor. The sense of smell detects very minute amounts of material and is usually much more sensitive in detecting the presence of volatile chemicals, but the tongue is able to detect infinitesimal amounts of some vapors if it is protruded from the mouth so that solution of the gases in the saliva may take place. In this manner traces of sulfur dioxide can be detected in the air since it dissolves in the saliva and creates a sour taste.

Flavors described as hot are those that exert a mild counterirritant effect on the mucosa of the mouth, those that are astringent and pucker the mouth contain tannins and acids that produce this effect by reacting with the lining of the mouth and wines possess a bouquet due to the odor of the volatile constituents. Indian turnip (Jack-in-the-pulpit) owes its flavor largely to the stinging sensation caused by the minute acicular crystals of calcium oxalate which penetrate the mucous membrane.

Other physiological and physical factors that also may affect taste are coarseness or grittiness due to small particles, eg, ion-exchange resins. Antidiarrheal preparations have a chalky taste. Menthol imparts a cool taste because it affects the coldness receptors. Mannitol gives a cool sensation when it dissolves because its negative of heat of solution will cause the temperature to drop. For this reason, mannitol often is used as the base for chewable tablets.

There is a definite threshold of taste for every substance, which varies somewhat with the individual and with the environment. The experienced chef tastes his delicacies at the temperature at which they will be served since heat and cold alter the flavor of many preparations. Thus, lemon

loses its sour taste entirely at an elevated temperature and other flavors become almost nonvolatile, tasteless and odorless when cooled sufficiently. In addition to the influence of temperature, the sensitivity of each individual must be considered. For example, it has been determined by experiment that the amount of sugar that can just be detected by the average individual is about 7 mg. However, this amount cannot be tasted by some and it is definitely sweet to others.

People are more sensitive to odor than to taste. There are about 10,000 to 30,000 identifiable scents, of which the average person can identify about 4000. Women are more sensitive to odors than men. Additional insights can be obtained by reading Cagan RH, Kare MR: *Biochemistry of Taste and Olfaction*, Academic, 1981, and Beidler LM (ed): *Handbook of Sensory Physiology*, vol IV, pts 1 and 2, Springer-Verlag, 1971.

Preservation of Flavors—Most monographs of official products contain specific directions for storage. Proper methods of storage are essential to prevent deterioration which in many instances results in destruction of odor and taste. Under adverse conditions undesirable changes occur due to one or a combination of the following: enzymatic activity, oxidation, change in moisture content, absorption of odors, activity of microorganisms and effects of heat and light. In certain products some of the changes wrought by these factors are desirable, as when esters are formed due to the activity of enzymes and when blending and mellowing results from the interchange of the radicals of esters (*transesterification*).

One method for protecting readily oxidizable substances, such as lemon oil, from deteriorating, and thus preserving their original delicate flavor, is to microencapsulate them by spray-drying. The capsules containing the flavors then are enclosed in various packaged products (eg, powdered gelatins) or tablets which are flavored deliciously when the capsule is disintegrated by mixing and warming with water or saliva.

Correlation of Chemical Structure with Flavor and Odor—The compounds employed as flavors in vehicles vary considerably in their chemical structure, ranging from simple esters (methyl salicylate), alcohols (glycerin) and aldehydes (vanillin) to carbohydrates (honey) and the complex volatile oils (anise oil). Synthetic flavors of almost any desired type are now available. These frequently possess the delicate flavor and aroma of the natural products and also the desirable characteristics of stability, reproducibility and comparatively low cost. Synthetic products such as cinnamaldehyde and benzaldehyde, first officially recognized when several of the essential oils became scarce during World War II, have been used widely.

There is a close relationship between chemical structure and taste. Solubility, the degree of ionization and the type of ions produced in the saliva definitely influence the sensation interpreted by the brain.

Sour taste is caused by hydrogen ions and it is proportional to the hydrogen-ion concentration and the lipid solubility of the compound. It is characteristic of acids, tannins, alum, phenols and lactones. Saltiness is due to simultaneous presence of anions and cations, eg, KBr, NH_4Cl and sodium salicylate. High-molecular-weight salts may have a bitter taste. Sweet taste is due to polyhydroxy compounds, polyhalogenated aliphatic compounds and α-amino acids. Amino and amide groups, especially if the positive effect is balanced by the proximity of a negative group, may produce a sweet taste. Sweetness increases with the number of hydroxy groups, possibly due to increase in solubility. Imides such as saccharin and sulfamates such as cyclamates are intensely sweet. Cyclamates have been removed from the market because they reportedly cause bladder tumors in rats. Free bases such as alkaloids and amides such as amphetamines give bitter tastes. Polyhydroxy compounds with a molecular weight greater than 300, halogenated substances and aliphatic thio compounds also may have bitter tastes. Unsaturation frequently bestows a sharp, biting odor and taste upon compounds.

No precise relationship between chemical structure and odor has been found. There are no primary odors, and odors blend into each other. Polymerization reduces or destroys odor; high valency gives odor and unsaturation enhances odor. A tertiary carbon atom often will give a camphoraceous odor, esters and lactones have a fruity odor and ketones have a pleasant odor. Strong odors often are accompanied by volatility and chemical reactivity.

Selection of Flavors

The proper selection of flavors for disguising nauseating medicines aids in their ingestion. Occasionally, sensitive patients have become nauseated sufficiently to vomit at the thought of having to take disagreeable medication, and it is particularly difficult to persuade children to continue to use and retain distasteful preparations. There is a need to know the allergies and idiosyncrasies of the patient; thus, it is foolish to use a chocolate-flavored vehicle for the patient who dislikes the flavor or who is allergic to it, notwithstanding the fact that this flavor is generally acceptable.

Flavoring Methodology

Each flavoring problem is unique and requires an individual solution. The problem of flavoring is further complicated because flavor and taste depend on individual preferences. In solving flavoring problems the following techniques have been used:

1. *Blending*—Fruit flavors blend with sour taste; bitter tastes can be blended with salty, sweet and sour tastes; salt reduces sourness and increases sweetness; chemicals such as vanillin, monosodium glutamate and benzaldehyde are used for blending.
2. *Overshadow*—Addition of a flavor whose intensity is longer and stronger than the obvious taste, eg, methyl salicylate, glycyrrhiza and oleoresins.
3. *Physical*—Formation of insoluble compounds of the offending drug, eg, sulfonamides; emulsification of oils; effervescence, eg, magnesium citrate solution; high viscosity of fluids to limit contact of drug with the tongue, and mechanical procedures such as coating tablets, are physical methods to reduce flavoring problems.
4. *Chemical*—Adsorption of the drug on a substrate, or formation of a complex of the drug with ion-exchange resins or complexing agents.
5. *Physiological*—The taste buds may be anesthetized by menthol or mint flavors.

Flavors, as used by the pharmacist in compounding prescriptions, may be divided into four main categories according to the type of taste which is to be masked, as follows:

1. *Salty Taste*—Cinnamon syrup has been found to be the best vehicle for ammonium chloride, and other salty drugs such as sodium salicylate and ferric ammonium citrate. In a study of the comparative efficiency of flavoring agents for disguising salty taste, the following additional vehicles were arranged in descending order of usefulness: orange syrup, citric acid syrup, cherry syrup, cocoa syrup, wild cherry syrup, raspberry syrup, glycyrrhiza elixir, aromatic elixir and glycyrrhiza syrup. The last-named is particularly useful as a vehicle for the salines by virtue of its colloidal properties and the sweetness of both glycyrrhizin and sucrose.
2. *Bitter Taste*—Cocoa syrup was found to be the best vehicle for disguising the bitter taste of quinine bisulfate, followed, in descending order of usefulness, by raspberry syrup, cocoa syrup, cherry syrup, cinnamon syrup, compound sarsaparilla syrup, citric acid syrup, licorice syrup, aromatic elixir, orange syrup and wild cherry syrup.
3. *Acrid or Sour Taste*—Raspberry syrup and other fruit syrups are especially efficient in masking the taste of sour substances such as hydrochloric acid. Acacia syrup and other mucilaginous vehicles are best for disguising the acrid taste of substances, such as capsicum, since they tend to form a colloidal protective coating over the taste buds of the tongue. Tragacanth, unlike acacia, may be used in an alcoholic vehicle.

4. *Oily Taste*—Castor oil may be made palatable by emulsifying with an equal volume of aromatic rhubarb syrup or with compound sarsaparilla syrup. Cod liver oil is disguised effectively by adding wintergreen oil or peppermint oil. Lemon, orange and anise or combinations of these are also useful. It is better to mix most of the flavor with the oil before emulsifying it, and then the small remaining quantity can be added after the primary emulsion is formed.

Those flavors that are most pleasing to the majority of people are associated with some stimulant of a physical or physiological nature. This may be a central nervous stimulant such as caffeine, which is the reason so many enjoy tea and coffee as a beverage, or it may be a counterirritant such as one of the spices that produce a "biting" sensation or an agent which "tickles" the throat such as soda water. Sherry owes its sharp flavor to its acetaldehyde content, and some of the volatile oils contain terpenes that are stimulating to the mucous surfaces.

Selection of Vehicles

Too few pharmacists realize the unique opportunity they have in acquainting physicians with a knowledge of how to increase both the palatability and efficacy of their prescribed medicines through the judicious selection of vehicles. Because of the training a pharmacist receives, his knowledge of the characteristics of various pharmaceuticals and therapeutic agents and his technique and skill in preparing elegant preparations are well-developed, so that he is qualified admirably to advise concerning the proper use of vehicles.

A large selection of flavors is available as well as a choice of colors, so that one may prescribe a basic drug for a prolonged period, but by changing the vehicle from time to time, the taste and appearance are so altered that the patient does not tire of the prescription or show other psychological reactions to it.

The statement of the late Dr Bernard Fantus that "the best solvent is the best vehicle" helps to explain the proper use of a flavoring vehicle. For example, a substance that is soluble in alcohol, eg, phenobarbital, will not leave an alcoholic vehicle readily to dissolve in the aqueous saliva.

Waters—These are the simplest of the vehicles and are available with several flavors. They contain no sucrose, a fact to be considered at times, since sucrose under certain circumstances may be undesirable. They are likewise nonalcoholic, another fact which frequently influences vehicle selection.

Elixirs—These have added sweetness that waters lack, and they usually contain alcohol, which imparts an added sharpness to the flavor of certain preparations, making the latter more pleasing to the taste. Elixirs are suitable for alcohol-soluble drugs.

Syrups—These vehicles, like elixirs, offer a wide selection of flavors and colors from which to choose. Their specific value, however, lies particularly in the fact that they are intensely sweet and contain little or no alcohol, a combination which makes them of singular value as masking agents for water-soluble drugs.

Vehicles consisting of a solution of pleasantly flavored volatile oils in syrup or glycerin (1:500) have been employed successfully in producing uniform and stable preparations. These vehicles are prepared by adding 2 mL of the volatile oil, diluted with 6 mL of alcohol, to 500 mL of glycerin or syrup, which has been warmed gently. The solution is added a little at a time with continuous shaking, and then sufficient glycerin or syrup is added to make 1000 mL, and mixed well.

Alcohol solutions of volatile oils are sometimes used as "stock solutions" for flavoring pharmaceuticals.

A listing of substances, most of them official, used as flavors, flavored vehicles or as sweeteners, is given in Table I. Additional information on flavoring ingredients may be obtained in Furia TE, Bellanca A: *Fenaroli's Handbook of Flavor Ingredients*, Chemical Rubber, Cleveland, 1971.

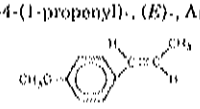
Table I—Flavoring Agents

Acacia syrup	Honey
Anethole	Iso-Alcoholic elixir
Anise oil	Lavender oil
Aromatic elixir	Lemon oil
Benzaldehyde	Lemon tincture
Benzaldehyde elixir, compound	Mannitol
	Methyl salicylate
Caraway	Nutmeg oil
Caraway oil	Orange, bitter, elixir
Cardamom oil	Orange, bitter, oil
Cardamom seed	Orange flower oil
Cardamom spirit, compound	Orange flower water
	Orange oil
Cardamom tincture, compound	Orange peel, bitter
	Orange peel, sweet, tincture
Cherry juice	Orange spirit, compound
Cherry syrup	Orange syrup
Cinnamon	Peppermint
Cinnamon oil	Peppermint oil
Cinnamon water	Peppermint spirit
Citric acid	Peppermint water
Citric acid syrup	Phenylethyl alcohol
Clove oil	Raspberry juice
Cocoa	Raspberry syrup
Cocoa syrup	Rosemary oil
Coriander oil	Rose oil
Dextrose	Rose water
Eriodictyon	Rose water, stronger
Eriodictyon fluidextract	Saccharin
Eriodictyon syrup, aromatic	Saccharin calcium
	Saccharin sodium
Ethyl acetate	Sarsaparilla syrup, compound
Ethyl vanillin	
Fennel oil	Sorbitol solution
Ginger	Spearmint
Ginger fluidextract	Spearmint oil
Ginger oleoresin	Sucrose
Glucose	Syrup
Glycerin	Thyme oil
Glycyrrhiza	Tolu balsam
Glycyrrhiza elixir	Tolu balsam syrup
Glycyrrhiza extract	Vanilla
Glycyrrhiza extract, pure	Vanilla tincture
Glycyrrhiza fluidextract	Vanillin
Glycyrrhiza syrup	Wild cherry syrup

Acacia Syrup—see page 1301.

Anethole

Benzene, 1-methoxy-4-(1-propenyl)-, (*E*)-, Anethol; Anise Camphor



(*E*)-*p*-Propenylanisole [4180-23-8] $C_{10}H_{12}O$ (148.20); obtained from anise oil and other sources, or prepared synthetically.

Preparation—It is the principal constituent of anise and fennel oil and usually is obtained from these sources by fractionating and chilling the proper fraction whereby it crystallizes out.

Description—Colorless or faintly yellow liquid at or above 23°; aromatic odor of anise and a sweet taste; affected by light; specific gravity 0.983 to 0.988; distils completely 231 to 237° and congeals at not less than 20°; its alcohol solution is neutral to litmus.

Solubility—Very slightly soluble in water; freely soluble in alcohol; miscible with chloroform or ether; yields a clear solution with 2 volumes of alcohol.

Uses—A *flavoring agent*. Its uses are similar to those of anise oil. It sometimes is sold as *Synthetic* or *Artificial Anise Oil* for flavoring and is a licorice-like flavor used in *Diphenhydramine Hydrochloride Elixir*.

Anise Oil

Aniseed Oil; Star Anise Oil

The volatile oil distilled with steam from the dried, ripe fruit of *Pimpinella anisum* Linné (Fam *Umbelliferae*) or from the dried, ripe fruit of *Illicium verum* Hooker filius (Fam *Magnoliaceae*).

Note—If solid material has separated, carefully warm the oil until it is completely liquefied, and mix it before using.

Constituents—The official oil varies somewhat in composition, depending upon whether it was obtained from *Pimpinella anisum* or the star anise, *Illicium verum*. *Anethole* is the chief constituent of both oils, occurring to the extent of 80 to 90%. *Methyl chavicol*, an isomer of anethole, and *anisic ketone* [$C_{10}H_{12}O_2$] are also found in both oils, as are small amounts of many other constituents.

Description—Colorless or pale yellow, strongly refractive liquid, having the characteristic odor and taste of anise; specific gravity 0.978 to 0.988; congeals not below 15°.

Solubility—Soluble in 3 volumes of 90% alcohol.

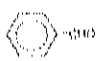
Uses—Extensively as a *flavoring agent*, particularly for licorice candies. It has been given as a *carminative* in a dose of about 0.1 mL.

Aromatic Elixir—page 1302.

Aromatic Elixir, Red—RPS-15, page 1240.

Benzaldehyde

Artificial Essential Almond Oil



Benzaldehyde {100-52-7} C_7H_6O (106.12).

Preparation—By the interaction of benzal chloride with lime in the presence of water. Benzal chloride is obtained by treating boiling toluene with chlorine.

Description—Colorless, strongly refractive liquid, having an odor resembling that of bitter almond oil, and a burning aromatic taste; affected by light; specific gravity 1.041 to 1.046; boils about 180°, solidifies about −56.5° and on exposure to air it gradually oxidizes to benzoic acid.

Solubility—Dissolves in about 350 volumes of water; miscible with alcohol, ether, chloroform or fixed and volatile oils.

Uses—In place of bitter almond oil for *flavoring* purposes; it is much safer than the latter because it contains no hydrocyanic acid. It also is used extensively in *perfumery* and in the manufacture of dyestuffs and many other organic compounds, such as aniline, acetanilid or mandelic acid.

Compound Benzaldehyde Elixir—*Preparation:* Dissolve benzaldehyde (0.5 mL) and vanillin (1 g) in alcohol (50 mL); add syrup (400 mL), orange flower water (150 mL) and sufficient purified water, in several portions, shaking the mixture thoroughly after each addition, to make the product measure 1000 mL; then filter, if necessary, until the product is clear. *Alcohol Content:* 3 to 5%. *Uses:* A useful vehicle for administering bromides and other salts, especially when a low alcoholic content is desired.

Camphor Water—RPS-13, page 436.

Caraway

Carum; Caraway Seed; Caraway Fruit; Kümmel

The dried ripe fruit of *Carum carvi* Linné (Fam *Umbelliferae*).

Constituents—About 5% of *volatile oil*, with a little *fixed oil* and other constituents.

Uses—A *flavor*. It also has been used empirically as a *carminative* and *stimulant*.

Caraway Oil [Oleum Cari]—A volatile oil distilled from the dried, ripe fruit of *Carum carvi* Linné (Fam *Umbelliferae*); yields not less than 50% (*v/v*) of $C_{10}H_{14}O$ (carvone). The chief odoriferous component of the oil is the ketone *d-carvone* [$C_{10}H_{14}O$], which is the optical isomer of the levorotatory variety occurring in spearmint oil. The remainder of the oil consists mainly of the terpene *d-limonene* [$C_{10}H_{16}$]. Colorless or pale yellow liquid, with the characteristic odor and taste of caraway; specific gravity 0.900 to 0.910. *Uses:* In making caraway water and as a flavor and *carminative* in other pharmaceutical preparations.

Cardamom Seed

Cardamom Fruit; Cardamom; Ceylon or Malabar Cardamom

The dried ripe seed of *Elettaria cardamomum* (Linné) Maton (Fam. *Zingiberaceae*).

It should be removed recently from the capsule.

Constituents—A *volatile oil*, the yield of which is 1.3% from Malabar Ceylon Seeds and 2.6% from Mysore-Ceylon Seeds. *Fixed oil* is present to the extent of 10%, also starch, mucilage, etc.

Uses—A *flavor*. For many years it was employed empirically as a *carminative*.

Cardamom Oil—The volatile oil distilled from the seed of *Elettaria cardamomum* (Linné) Maton (Fam *Zingiberaceae*). Varieties of the oil contain *d-α-terpineol* [$C_{10}H_{17}OH$] both free and as the acetate, 5 to 10% *cineol* [$C_{10}H_{18}O$] and *limonene* [$C_{10}H_{16}$]. The Ceylon Oil, however, contains the alcohol *4-terpineol* (*4-carbamenthenol*) [$C_{10}H_{17}OH$], the terpenes *terpinene* and *sabinene*, and *acetic* and *formic acids*, probably combined as esters. Colorless or very pale yellow liquid possessing the aromatic, penetrating and somewhat camphoraceous odor of cardamom, and a persistently pungent, strongly aromatic taste; affected by light. Specific gravity 0.917 to 0.947; miscible with alcohol; dissolves in 5 volumes of 70% alcohol. *Uses:* A *flavor*.

Cardamom Tincture, Compound—page 1302.

Cherry Juice—page 1320.

Cherry Syrup—page 1301.

Cinnamon

Saigon Cinnamon; True Cinnamon; Saigon Cassia

The dried bark of *Cinnamomum loureirii* Nees (Fam. *Lauraceae*).

It contains, in each 100 g, not less than 2.5 mL of volatile oil.

Uses—A *flavoring agent*. Formerly, it was used as a carminative.

Cinnamon Oil [Cassia Oil; Oil of Chinese Cinnamon]—The volatile oil distilled with steam from the leaves and twigs of *Cinnamomum cassia* (Nees) Nees ex Blume (Fam *Lauraceae*), rectified by distillation; contains not less than 80%, by volume, of the total aldehydes of cinnamon oil. Cinnamaldehyde is the chief constituent. Yellowish or brownish liquid, becoming darker and thicker on aging or exposure to the air, and having the characteristic odor and taste of cassia cinnamon; specific gravity 1.045 to 1.063. Soluble in an equal volume of alcohol, 2 volumes of 70% alcohol or an equal volume of glacial acetic acid. *Uses:* A *flavor*. It formerly was used in a dose of 0.1 mL for flatulent colic.

Cocoa

Cacao USP XVI; Prepared Cocoa; Powdered Cocoa; Cocoa Powder; Medium-Fat Cocoa

A powder prepared from the roasted, cured kernels of the ripe seed of *Theobroma cacao* Linné (Fam *Sterculiaceae*).

It yields 10 to 22% of nonvolatile, ether-soluble extractive.

Preparation—The cocoa bean is dark as the result of a fermentation and roasting process which it undergoes. *Plain chocolate* consists of shelled cocoa beans (*cocoa nibs*) ground to a smooth paste which forms a hard cake when it cools because of the high fat content (50 to 58%).

It is the food prepared by pulverizing the residue remaining after part of the fat has been removed by expression from plain chocolate. It may be flavored by the addition of ground spice, ground vanilla bean, vanillin, ethylvanillin, coumarin, salt and other flavors as long as they do not imitate the flavor of chocolate, milk or butter. Three types are recognized depending on fat content: *breakfast cocoa* or *high fat cocoa* (22% minimum), *cocoa* or *medium-fat cocoa* (10 to 22%) and *low-fat cocoa* (less than 10%).

Sweet chocolate is plain chocolate plus added sugar and flavor (usually vanilla).

Milk chocolate is a mixture of sweet chocolate and milk powder or other dairy product. Chocolate and the products described above contain the purines theobromine and caffeine, and considerable quantities of fat (cocoa butter or theobroma oil), as well as protein and starch. These factors are lowered in sweet chocolate because of the large amount of added sugar (more than 50% of the final product).

Description—Weak reddish to purplish brown to moderate brown powder having a chocolate-like odor and taste, free from sweetness.

Uses—A food and pharmaceutically as a flavor in tablets, syrups, pill and tablet coatings, troches, etc.

Cocoa Syrup—page 1301.

Coriander—page 1200.

Coriander Oil

The volatile oil distilled with steam from the dried ripe fruit of *Coriandrum sativum* Linné (Fam *Umbelliferae*).

Constituents—The alcohol *d-linalool* (formerly termed *"coriandrol"*) is the chief constituent of this oil, occurring in amounts varying from 60 to 80%. Other constituents include *l-borneol*, *geraniol*, *pinenes*, *terpinenes* and *p-cymene*.

Description—Colorless or pale yellow liquid, having the characteristic odor and taste of coriander; specific gravity 0.863 to 0.875.
Solubility—Soluble in 3 volumes of 70% alcohol.

Uses—A flavoring agent. It formerly was employed in a dose of 0.1 mL as a *carminative*.

Denatonium Benzoate—page 1321.

Eriodictyon

Consumptives' Weed; Mountain Balm; Yerba Santa

The dried leaf of *Eriodictyon californicum* (Hooker et Arnott) Torrey (Fam *Hydrophyllaceae*).

Constituents—A bitter *resin*, *volatile oil*, *eriodictyonone* [$C_{16}H_{14}O_6$, also called *homoeriodictyol*], *fixed oil*, *tannin*, *gum*, etc.
Uses—A pharmaceutical necessity. It is used in the preparation of *Eriodictyon Fluidextract*.

Eriodictyon Fluidextract [Yerba Santa Fluidextract]—*Preparation:* Using Eriodictyon (in moderately coarse powder, 1000 g), prepare the fluidextract by Process A (page 1543), using a mixture of 4 volumes of alcohol and 1 volume of water as the menstruum. Macerate the drug during 48 hr, then percolate at a moderate rate and reserve the first 800 mL of percolate. *Alcohol Content:* 57 to 62%. *Uses:* A peculiar, aromatic *flavor* used in syrups and elixirs, especially for masking the taste of bitter drugs like quinine. Because of its resinous character it requires an alkali to render it soluble in aqueous mixtures.

Eriodictyon Syrup, Aromatic—page 1301.

Ethyl Acetate

Acetic acid, ethyl ester; Acetic Ether

$CH_3COOC_2H_5$

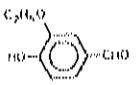
Ethyl acetate [141-78-6] $C_4H_8O_2$ (88.11).
Preparation—By slow distillation of a mixture of alcohol and acetic acid in the presence of sulfuric acid.

Description—Transparent, colorless liquid with a fragrant and refreshing, slightly acetous odor, and a peculiar acetous, burning taste; specific gravity 0.894 to 0.898; distills 76 to 77.5°.
Solubility—1 mL in about 10 mL of water; miscible with alcohol, acetone, ether, chloroform or fixed and volatile oils.

Uses—Chiefly as a *flavoring agent*. It is used industrially in artificial fruit essence, as *a solvent* for nitrocellulose varnishes and lacquers and as a solvent in organic chemistry.

Ethyl Vanillin

Benzaldehyde, 3-ethoxy-4-hydroxy-; Bourbonal; Ethovan; Vanillal; Vanirome



3-Ethoxy-4-hydroxybenzaldehyde [121-32-4] $C_9H_{10}O_3$ (166.18).
Preparation—By reacting *o*-ethoxyphenol with formaldehyde and *p*-nitrosodimethylaniline in the presence of aluminum and water.

Description—Fine, white or slightly yellowish crystals; odor and taste similar to vanillin; affected by light; solutions are acid to litmus; melts about 77°.
Solubility—1 g in about 100 mL of water at 50°; freely soluble in alcohol, chloroform, ether or solutions of fixed alkali hydroxides.

Uses—A *flavor*, like vanillin, but stronger.

Eucalyptus Oil

The volatile oil distilled with steam from the fresh leaf of *Eucalyptus globulus* Labillardière or of some other species of *Eucalyptus* L'Heritier (Fam *Myrtaceae*). It contains not less than 70% of $C_{10}H_{18}O$ (eucalyptol).

Constituents—The most important constituent is *eucalyptol* (*cineol*). Other compounds include *d-α-pinene*, *globulol*, *pinocarveol*, *pinocarvone* and several aldehydes.

Description—Colorless or pale yellow liquid, having a characteristic, aromatic, somewhat camphoraceous odor, and a pungent, spicy, cooling taste; specific gravity 0.905 to 0.925 at 25°.
Solubility—Soluble in 5 volumes of 70% alcohol.

Uses—A *flavoring agent* and an *expectorant* in chronic bronchitis. It also has *bacteriostatic* properties. This oil may be toxic.

Fennel Oil

The volatile oil distilled with steam from the dried ripe fruit of *Foeniculum vulgare* Miller (Fam *Umbelliferae*).

Note—If solid material has separated, carefully warm the oil until it is completely liquefied, and mix it before using.

Constituents—*Anethole* [$C_{10}H_{12}O$] is the chief constituent, occurring to the extent of 50 to 60%. Some of the other constituents are *d*-pinene, phellandrene, dipentene, fenchone, methylchavicol, anisaldehyde and anisic acid.

Description—Colorless or pale yellow liquid, having the characteristic odor and taste of fennel; specific gravity 0.953 to 0.973; congealing temperature is not below 3°.
Solubility—Soluble in 3 volumes of 80% alcohol or in 1 volume of 90% alcohol.

Uses—A flavoring agent. It formerly was employed in a dose of 0.1 mL as a *carminative*.

Glycyrrhiza

Licorice Root; Liquorice Root; Sweetwood; Italian Juice Root; Spanish Juice Root

The dried rhizome and roots of *Glycyrrhiza glabra* Linné, known in commerce as Spanish Licorice, or of *Glycyrrhiza glabra* Linné var *glandulifera* Waldstein et Kitaibel, known in commerce as Russian Licorice, or of other varieties of *Glycyrrhiza glabra* Linné, yielding a yellow and sweet wood (Fam. *Leguminosae*).

Constituents—This well-known root contains 5 to 7% of the sweet principle *glycyrrhizin*, or *glycyrrhizic acid* which is 50 times as sweet as cane sugar. There also is present an oleoresinous substance to which its slight acridity is due. If alcohol or an alkali is used as a menstruum for the root and the preparation not treated to deprive it of acridity, it will have a disagreeable aftertaste. For this reason boiling water is used for its extraction in both the extract and the fluidextract.

Description—The USP/NF provides descriptions of *Unground Spanish and Russian Glycyrrhizas*, *Histology* and *Powdered Glycyrrhiza*.

Uses—Valuable in pharmacy chiefly for its *sweet flavor*. It is one of the most efficient substances known for masking the taste of bitter substances, like quinine. Acids precipitate the glycyrrhizin and should not be added to mixtures in which glycyrrhiza is intended to mask disagreeable taste. Most of the imported licorice is used

by tobacco manufacturers to flavor tobacco. It also is used in making candy.

Pure Glycyrrhiza Extract [Pure Licorice Root Extract]—*Preparation:* Moisten 1000 g of glycyrrhiza, in granular powder, with boiling water, transfer it to a percolator, and percolate with boiling water until the glycyrrhiza is exhausted. Add enough diluted ammonia solution to the percolate to impart a distinctly ammoniacal odor, then boil the liquid under normal atmospheric pressure until it is reduced to a volume of about 1500 mL. Filter the liquid, and immediately evaporate the filtrate until the residue has a pilular consistency. Pure extract of glycyrrhiza differs from the commercial extract in that it is almost completely soluble in aqueous mixtures. The large amount of filler used in the commercial extract to give it firmness renders it unfit to use as a substitute for the pure extract. *Description:* Black, pilular mass having a characteristic, sweet taste. *Uses:* A *flavoring agent*. One of the ingredients in *Aromatic Cascara Sagrada Fluidextract*.

Glycyrrhiza Fluidextract [Licorice Root Fluidextract; Liquid Extract of Liquorice]—*Preparation:* To 1000 g of coarsely ground glycyrrhiza add about 3000 mL of boiling water, mix, and allow to macerate in a suitable, covered percolator for 2 hr. Then allow the percolation to proceed at a rate of 1 to 3 mL/min, gradually adding boiling water until the glycyrrhiza is exhausted. Add enough diluted ammonia solution to the percolate to impart a distinctly ammoniacal odor, then boil the liquid actively under normal atmospheric pressure until it is reduced to a volume of about 1500 mL. Filter the liquid, evaporate the filtrate on a steam bath until the residue measures 750 mL, cool, gradually add 250 mL of alcohol and enough water to make the product measure 1000 mL, and mix. *Alcohol Content:* 20 to 24%, by volume. *Uses:* A pleasant *flavor* for use in syrups and elixirs to be employed as vehicles and correctives.

Glycyrrhiza Elixir—page 1302.
Glycyrrhiza Syrup—page 1302.
Honey—page 1302.
Hydriodic Acid Syrup—page 1302.
Iso-Alcoholic Elixir—page 1328.

Lavender Oil

Lavender Flowers Oil

The volatile oil distilled with steam from the fresh flowering tops of *Lavandula officinalis* Chaix ex Villars (*Lavandula vera* DeCandolle) (Fam *Labiatae*) or produced synthetically. It contains not less than 35% of esters calculated as $C_{12}H_{20}O_2$ (linalyl acetate).

Constituents—It is a product of considerable importance in perfumery. *Linalyl acetate* is the chief constituent. *Cineol* appears to be a normal constituent of English oils. Other constituents include *amyl alcohol*, *d-borneol* (small amount); *geraniol*, *lavandulal* ($C_{10}H_{16}O$); *linalool*; *nerol*; *acetic*, *butyric*, *valeric*, and *caproic acids* (as esters); traces of *d-pinene*, *limonene* (in English oils only) and the sesquiterpene *caryophyllene*; *ethyl n-amyl ketone*; an aldehyde (probably *valeric aldehyde*) and *coumarin*.

Description—Colorless or yellow liquid, having the characteristic odor and taste of lavender flowers; specific gravity 0.875 to 0.888.

Solubility—1 volume dissolves in 4 volumes of 70% alcohol.

Uses—Primarily as a *perfume*. It formerly was used in doses of 0.1 mL as a *carminative*.

Lemon Oil

The volatile oil obtained by expression, without the aid of heat, from the fresh peel of the fruit of *Citrus limon* (Linné) Burmann filius (Fam *Rutaceae*), with or without the previous separation of the pulp and the peel. The total aldehyde content, calculated as citral ($C_{10}H_{16}O$), is 2.2–3.8% for California-type oil, and 3.0–5.5% for Italian-type oil.

Note—Do not use oil that has a terebinthine odor.

Constituents—From the standpoint of odor and flavor, the most noteworthy constituent is the aldehyde *citral*, which is present to the extent of about 4%. About 90% of *d-limonene* is present; small amounts of *l-α-pinene*, *β-pinene*, *camphene*, *β-phellandrene* and γ-terpinene also occur. About 2% of a solid, nonvolatile substance called *citroptene*, *limettin* or *lemon-camphor*, which is dissolved out of the peel, also is present. In addition, there are traces of several other compounds: *α-terpineol*; the *acetates of linalool* and *geranial*; *citronellal*, *octyl* and *nonyl aldehydes*; the sesquiterpenes *bisabolene* and *cadinene* and the ketone *methylheptenone*.

When fresh, the oil has the fragrant odor of lemons. Because of the instability of the terpenes present, the oil readily undergoes deterioration by oxidation, acquiring a terebinthinate odor.

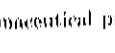
Description—Pale yellow to deep yellow or greenish yellow liquid, with the characteristic odor and taste of the outer part of fresh lemon peel; specific gravity 0.849 to 0.955.

Solubility—Soluble in 3 volumes of alcohol; miscible in all proportions with dehydrated alcohol, carbon disulfide or glacial acetic acid.

Uses—A *flavor* in pharmaceutical preparations and in certain candies and foods.

Methyl Salicylate

Benzoic acid, 2-hydroxy-, methyl ester; Gaultheria Oil; Wintergreen Oil; Betula Oil; Sweet Birch Oil; Teaberry Oil; Artificial Wintergreen Oil; Synthetic Wintergreen Oil



Methyl salicylate [119-36-8] $C_6H_4(OH)COOCH_3$ (152.15); produced synthetically or obtained by maceration and subsequent distillation with steam from the leaves of *Gaultheria procumbens* Linné (Fam *Ericaceae*) or from the bark of *Betula lenta* Linné (Fam *Betulaceae*).

Note—It must be labeled to indicate whether it was made synthetically or distilled from either of the plants mentioned above.

Preparation—Found naturally in gaultheria and betula oils and in many other plants but the commercial product is usually synthetic, made by esterifying salicylic acid with methyl alcohol in the presence of sulfuric acid and distilling.

Description—Colorless, yellowish or reddish liquid, having the characteristic odor and taste of wintergreen; specific gravity (synthetic), 1.180 to 1.185, (from gaultheria or betula), 1.176 to 1.182; boils between 219 to 224° with some decomposition.

Solubility—Slightly soluble in water; soluble in alcohol or glacial acetic acid.

Uses—A pharmaceutical necessity and *counterirritant* (local analgesic). As a pharmaceutical necessity, it is used to flavor the official *Aromatic Cascara Sagrada Fluidextract*, and it is equal in every respect to wintergreen oil or sweet birch oil. As a counterirritant, it is applied to the skin in the form of a liniment, ointment or cream; care should be exercised since salicylate is absorbed through the skin.

Caution—Because it smells like wintergreen candy, it is ingested frequently by children and has caused many fatalities. *Keep out of the reach of children.*

Dose—*Topical*, in lotions and solutions in 10 to 25% concentration.

Monosodium Glutamate

Glutamic acid, monosodium salt, monohydrate

[142-47-2] $C_5H_8NNaO_4 \cdot H_2O$ (187.13)

Preparation—From the fermentation of beet sugar or molasses or by hydrolysis of vegetable proteins.

Description—White, crystalline powder. The pentahydrate effloresces in air to form the monohydrate.

Solubility—Very soluble in water; sparingly soluble in alcohol.

Uses—Flavoring agent and perfume.

Nutmeg Oil

Myristica Oil NF XIII; East Indian Nutmeg Oil; West Indian Nutmeg Oil

The volatile oil distilled with steam from the dried kernels of the ripe seeds of *Myristica fragrans* Houttuyn (Fam *Myristicaceae*).

Constituents—It contains about 80% of *d-pinene* and *d-camphene*, 8% of *dipentene*, about 6% of the alcohols *d-borneol*, *geraniol*, *d-linalool* and *terpineol*, 4% of *myristicin*, 0.6% of *safrol*, 0.3% of *myristic acid* free and as esters, 0.2% of *eugenol* and *isoeugenol* and traces of the alcohol *terpineol-4*, a citral-like aldehyde and several acids, all present as esters.

Description—Colorless or pale yellow liquid having the characteristic odor and taste of nutmeg; specific gravity (East Indian Oil) 0.880 to 0.910, (West Indian Oil) 0.864 to 0.880.

Solubility—Soluble in an equal amount of alcohol; 1 volume of East Indian Oil in 3 volumes of 90% alcohol; 1 volume of West Indian Oil in 4 volumes of 90% alcohol.

Uses—Primarily as a *flavoring agent*. It is used for this purpose in *Aromatic Ammonia Spirit* (page 1533). The oil also is employed as a *flavor* in foods, certain alcoholic beverages, dentifrices and tobacco; to some extent, it also is used in perfumery. It *formerly* was used as a *carminative* and *local stimulant* to the gastrointestinal tract in a dose of 0.03 mL. In overdoses, it acts as a narcotic poison. *This oil is very difficult to keep and even if slightly terebinthinate is unfit for flavoring purposes.*

Orange Oil

Sweet Orange Oil

The volatile oil obtained by expression from the fresh peel of the ripe fruit of *Citrus sinensis* (Linné) Osbeck (Fam *Rutaceae*). The total aldehyde content, calculated as decanal ($C_{10}H_{20}O$), is 1.2 to 2.5%.

Note—Do not use oil that has a terebinthine odor.

Constituents—Consists of *d-limonene* to the extent of at least 90%; in the remaining 5 to 10% are the odorous constituents, among which, in samples of American origin, are *n-decylic aldehyde, citral, d-linalool, n-nonyl alcohol* and traces of *esters of formic, acetic, caprylic* and *capric* acids.

In addition to most of these compounds, Italian-produced oil contains *d-terpineol, terpinolene, α-terpinene* and *methyl anthranilate.*

Kept under the usual conditions it is very prone to decompose, and rapidly acquires a terebinthine odor.

Description—Intensely yellow orange or deep orange liquid, which possesses the characteristic odor and taste of the outer part of fresh sweet orange peel; specific gravity 0.842 to 0.846.

Solubility—Miscible with dehydrated alcohol and with carbon disulfide; dissolves in an equal volume of glacial acetic acid.

Uses—A *flavoring agent* in elixirs and other preparations.

Orange Flower Oil

Neroli Oil

The volatile oil distilled from the fresh flowers of *Citrus aurantium* Linné (Fam *Rutaceae*).

Constituents—*β-Ocimene, l-α-pinene, l-camphene, dipentene, l-linalool, geraniol, farnesol, d-terpineol, phenylethyl alcohol, nerol, nerolidol, decylic aldehyde, jasmone, methyl anthranilate, indole, acetic esters of the alcohols* present and traces of *esters of benzoic, phenylacetic* and *palmitic acids.*

Description—Pale yellow, slightly fluorescent liquid, which becomes reddish brown on exposure to light and air; distinctive, fragrant odor, similar to that of orange blossoms, and an aromatic, at first sweet, then somewhat bitter, taste; may become turbid or solid at low temperatures; specific gravity 0.863 to 0.880; neutral to litmus paper; an alcoholic solution has a *violet fluorescence.*

Uses—A *flavor* and *perfume*. Several less valuable varieties of the oil are known commercially. These are designated as *Bigarade* (from the fresh flowers of bitter orange, the ordinary neroli oil), *Portugal* (from the fresh flowers of sweet orange) and *Petit-grain* (from the leaves and young shoots of the bitter orange). The finest variety is known as *Petale.*

Orange Flower Water—page 1300.

Sweet Orange Peel Tincture

Preparation—From sweet orange peel, which is the outer rind of the nonartificially colored, fresh, ripe fruit of *Citrus sinensis* (Linné) Osbeck (Fam *Rutaceae*), by Process M (page 1543). Macerate 500 g of the sweet orange peel (*Note—Exclude the inner, white portion of the rind*) in 900 mL of alcohol, and complete the preparation with alcohol to make the product measure 1000 mL. Use talc as the filtering medium.

The white portion of the rind must not be used, as the proportion of oil, which is only in the yellow rind, is reduced, and the bitter principle *hesperidin* is introduced.

Alcohol Content—62 to 72%.

Uses—A *flavor*, used in syrups, elixirs and emulsions. This tincture was introduced to provide a delicate orange flavor direct from the fruit instead of depending upon orange oil which so frequently is terebinthinate and unfit for use. The tincture keeps well.

Compound Orange Spirit

Contains, in each 100 mL, 25 to 30 mL of the mixed oils.

Orange Oil	200 mL
Lemon Oil	50 mL
Coriander Oil	20 mL
Anise Oil	5 mL
Alcohol, a sufficient quantity,	
To make	1000 mL

Mix the oils with sufficient alcohol to make the product measure 1000 mL.

Alcohol Content—65 to 75%.

Uses—A *flavor* for elixirs. An alcoholic solution of this kind permits the uniform introduction of small proportions of oils and also preserves orange and lemon oils from rapid oxidation. These two oils should be bought in small quantities by the pharmacist, since the spirit is made most satisfactorily from oils taken from bottles not previously opened. This will insure that delicacy of flavor which should always be characteristic of elixirs.

Orange Syrup

Syrup of Orange Peel

Contains, in each 100 mL, 450 to 550 mg of citric acid ($C_6H_8O_7$).

Sweet Orange Peel Tincture	50 mL
Citric Acid (anhydrous)	5 g
Talc	15 g
Sucrose	820 g
Purified Water, a sufficient quantity,	
To make	1000 mL

Triturate the talc with the tincture and citric acid, and gradually add 400 mL of purified water. Then filter, returning the first portions of the filtrate until it becomes clear, and wash the mortar and filter with enough purified water to make the filtrate measure 450 mL. Dissolve the sucrose in this filtrate by agitation, without heating, and add enough purified water to make the product measure 1000 mL. Mix and strain.

Note—Do not use syrup that has a terebinthine odor or taste or shows other indications of deterioration.

Alcohol Content—2 to 5%.

Uses—A pleasant, acidic vehicle.

Peppermint

American Mint; Lamb Mint; Brandy Mint

Consists of the dried leaf and flowering top of *Mentha piperita* Linné (Fam *Labiatae*).

Uses—The source of green color for *Peppermint Spirit* (page 798). The odor of fresh peppermint is due to the presence of about 2% of a volatile oil, much of which is lost on drying the leaves in air. It is cultivated widely both in the US and France. It formerly was used as a carminative.

Peppermint Oil—The volatile oil distilled with steam from the fresh overground parts of the flowering plant of *Mentha piperita* Linné (Fam *Labiatae*), rectified by distillation and neither partially nor wholly dementhollized. It yields not less than 5% of esters, calculated as menthyl acetate ($C_{12}H_{22}O_2$), and not less than 50% of total menthol ($C_{10}H_{20}O$), free and as esters. *Constituents:* This is one of the most important of the group of volatile oils. The chief constituent is *Menthol* (page 765) which occurs in the levorotatory form; its ester, *menthyl acetate*, is present in a much smaller amount. Other compounds which are present include the ketone *menthone, piperitone, α-pinene, l-limonene, phellandrene, cadinene, menthyl isovalerate, isovaleric aldehyde, acetaldehyde, menthofuran, cineol,* an unidentified *lactone* ($C_{10}H_{16}O_2$) and probably *amyl acetate*. Colorless or pale yellow liquid, having a strong,

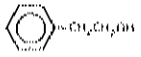
penetrating odor of peppermint and a pungent taste, followed by a sensation of cold when air is drawn into the mouth; specific gravity 0.896 to 0.908; 1 volume dissolves in 3 volumes of 70% alcohol. *Uses*—A *flavoring agent, carminative, antiseptic* and *local anesthetic*. It also is used extensively as a *flavor* in candy, chewing gum, etc.

Peppermint Spirit—page 798.

Peppermint Water—page 1300.

Phenylethyl Alcohol

Benzeneethanol; 2-Phenylethanol



Phenethyl alcohol [60-12-8] $C_8H_{10}O$ (122.17); occurs in a number of essential oils such as those of rose, neroli, hyacinth, carnation and others.

Description—Colorless liquid with a rose-like odor and a sharp, burning taste; solidifies at −27°; specific gravity 1.017 to 1.020.

Solubility—1 g in 80 mL of water; ≥1 mL of alcohol, chloroform or ether; very soluble in fixed oils, glycerin or propylene glycol; slightly soluble in mineral oil.

Uses—Introduced for use as an antibacterial agent in ophthalmic solutions, but it is of limited effectiveness.

It is used in *flavors*, as a *soap perfume* and in the preparation of synthetic oils of rose and similar flower oils. It is also a valuable perfume fixative.

Pine Needle Oil

Dwarf Pine Oil

The volatile oil distilled with steam from the fresh leaf of *Pinus mugo* Turra and its variety *pumilio* (Haenke) Zenari (Fam *Pinaceae*); contains 3 to 10%, by weight, of esters calculated as $C_{12}H_{20}O_2$ (bornyl acetate).

Constituents—It contains the terpenes *l-α-pinene, β-pinene, l-phellandrene, l-limonene, dipentene*, and possibly *sylvestrene*, the ester *bornyl acetate* and several unidentified terpene and sesquiterpene alcohols.

Description—Colorless to yellowish liquid, having a pleasant, aromatic odor and a bitter, pungent taste; specific gravity 0.853 to 0.871 at 25°.

Solubility—Dissolves in 4.5 to 10 volumes of 90% alcohol, often with turbidity.

Uses—Chiefly as a *perfume* and *flavoring agent*. It also is employed as an inhalant in bronchitis.

Raspberry Syrup—page 1302.

Rose Oil

Otto of Rose; Attar of Rose

The volatile oil distilled with steam from the fresh flowers of *Rosa gallica* Linné, *Rosa damascena* Miller, *Rosa alba* Linné, *Rosa centifolia* Linné and varieties of these species (Fam *Rosaceae*).

Constituents—From the quantitative standpoint the chief components are the alcohols *geraniol* [$C_{10}H_{18}O$] and *l-citronellol* [$C_{10}H_{20}O$]. The sesquiterpene alcohols *farnesol* and *nerol* occur to the extent of 1% and 5 to 10%, respectively. Together, the four alcohols constitute 70 to 75% of the oil. *Phenylethyl alcohol*, which comprises 1% of the oil, is an important odoriferous constituent. Other compounds present are *linalool, eugenol, nonyl aldehyde*, traces of *citral* and two solid hydrocarbons of the paraffin series.

Description—A colorless or yellow liquid, which has the characteristic odor and taste of rose; at 25°, a viscous liquid; on gradual cooling it changes to a translucent, crystalline mass, which may be liquefied easily by warming; specific gravity 0.848 to 0.863 at 30° compared with water at 15°; 1 mL mixes with 1 mL of chloroform without turbidity; on the addition of 20 mL of 90% alcohol to this solution, the resulting liquid is neutral or acid to moistened litmus paper and deposits a crystalline residue within 5 min on standing at 20°.

Uses—Principally as a *perfume*. It is recognized officially for its use as an ingredient in *Rose Water Ointment* and cosmetics.

Stronger Rose Water

Triple Rose Water

A saturated solution of the odoriferous principles of the flowers of *Rosa centifolia* Linné (Fam *Rosaceae*), prepared by distilling the fresh flowers with water and separating the excess volatile oil from the clear, water portion of the distillate.

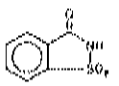
Note—When diluted with an equal volume of purified water, it may be supplied when *Rose Water* is required.

Description—Nearly colorless and clear liquid which possesses the pleasant odor and taste of fresh rose blossoms; must be free from empyreuma, mustiness and fungal growths.

Uses—An ingredient in *Rose Water Ointment*. It sometimes is prepared extemporaneously from concentrates or from rose oil, but such water is not official and rarely compares favorably with the fresh distillate from rose petals.

Saccharin

1,2-Benzisothiazol-3(2*H*)-one, 1,1-dioxide; Gluside; *o*-Benzosulfimide
Saxin (*Burroughs Wellcome*); Sweeta (*Squibb*)



1,2-Benzisothiazolin-3-one 1,1-dioxide [81-07-2] $C_7H_5NO_3S$ (183.18).

Preparation—Toluene is reacted with chlorosulfonic acid to form *o*-toluenesulfonyl chloride, which is converted to the sulfonamide with ammonia. The methyl group then is oxidized with dichromate yielding *o*-sulfamoylbenzoic acid which, when heated, forms the cyclic imide.

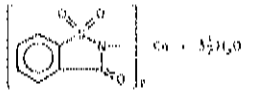
Description—White crystals or a white crystalline powder; odorless or has a faint aromatic odor; in dilute solution it is intensely sweet; solutions are acid to litmus; melts between 226 to 230°.

Solubility—1 g in 290 mL of water, 31 mL of alcohol or 25 mL of boiling water; slightly soluble in chloroform or ether; readily dissolved by dilute solution of ammonia, solutions of alkali hydroxides or solutions of alkali carbonates with the evolution of CO_2.

Uses—A sweetening agent in *Aromatic Cascara Sagrada Fluidextract* and highly alcoholic preparations. It is an intensely sweet substance. A 60-mg portion is equivalent in sweetening power to approximately 30 g of sucrose. It is used as a *sweetening agent* in vehicles, canned foods, beverages and in diets for diabetics to replace the sucrose. The relative sweetening power of saccharin is increased by dilution.

Saccharin Calcium

1,2-Benzisothiazol-3(2*H*)-one, 1,1-dioxide, calcium salt, hydrate (2:7)
Calcium *o*-Benzosulfimide



1,2-Benzisothiazolin-3-one 1,1-dioxide calcium salt hydrate (2:7) [6381-91-5] $C_{14}H_8CaN_2O_6S_2\cdot3\frac{1}{2}H_2O$ (467.48); *anhydrous* [6485-34-3] (404.43).

Preparation—Saccharin is reacted with a semimolar quantity of calcium hydroxide in aqueous medium and the resulting solution is concentrated to crystallization.

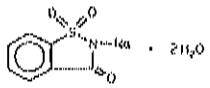
Description—White crystals or a white, crystalline powder; odorless or has a faint aromatic odor; and an intensely sweet taste even in dilute solutions; in dilute solution it is about 300 times as sweet as sucrose.

Solubility—1 g in 2.6 mL of water or 4.7 mL of alcohol.

Uses and Dose—See *Saccharin*.

Saccharin Sodium

1,2-Benzisothiazol-3(2*H*)-one, 1,1-dioxide, sodium salt, dihydrate; *Soluble Saccharin, Soluble Gluside; Sodium o-Benzosulfimide*



1,2-Benzisothiazolin-3-one 1,1-dioxide sodium salt dihydrate [6155-57-3] $C_7H_4NNaO_3S.2H_2O$ (241.19); *anhydrous* [128-44-9] (205.16).

Preparation—Saccharin is dissolved in an equimolar quantity of aqueous sodium hydroxide and the solution is concentrated to crystallization.

Description—White crystals or a white crystalline powder; odorless or has a faint aromatic odor and an intensely sweet taste even in dilute solutions; in dilute solution it is about 300 times as sweet as sucrose; when in powdered form it usually contains about ⅛ the theoretical amount of water of hydration due to efflorescence.

Solubility—1 g in 1.5 mL of water or 50 mL of alcohol.

Uses—Same as *Saccharin* but has the advantage of being more soluble in neutral aqueous solutions.

Application—15 to 60 mg as necessary.

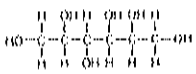
Dosage Form—Tablets: 15, 30 and 60 mg.

Sarsaparilla Syrup, Compound—RPS-13, page 445.

Sherry Wine—page RPS-15, page 1240.

Sorbitol

Sionin; Sorbit; D-Sorbitol; D-Glucitol Sorbo (*Atlas*)



D-Glucitol [50-70-4] $C_6H_{14}O_6$ (182.17); it may contain small amounts of other polyhydric alcohols.

Preparation—Commercially by reduction (hydrogenation) of certain sugars, such as glucose.

Description—White, hygroscopic powder, granules or flakes, having a sweet taste; the usual form melts about 96°.

Solubility—1 g in about 0.45 mL of water; slightly soluble in alcohol, methanol or acetic acid.

Uses—An *osmotic diuretic* given intravenously in 50% (*w/v*) solution to diminish edema, lower cerebrospinal pressure or reduce intraocular pressure in glaucoma. It also is used as a laxative, sweetener, humectant, plasticizer and, in 70% (*w/w*) solution, as a vehicle.

Dose—50 to 100 mL of a 50% solution; *laxative, oral*, 30 to 50 g.

Sorbitol Solution is a water solution containing, in each 100 g, 69–71 g of total solids consisting essentially of D-sorbitol and a small amount of mannitol and other isomeric polyhydric alcohols. The content of D-sorbitol [$C_6H_8(OH)_6$] in each 100 g is not less than 64 g. *Description:* Clear, colorless, syrupy liquid, having a sweet taste and no characteristic odor; neutral to litmus; specific gravity not less than 1.285; refractive index at 20° 1.455 to 1.465. *Uses:* It is not to be injected. It has been used as a replacement for propylene glycol and glycerin.

Spearmint

Spearmint Leaves; Spearmint Herb; Mint

The dried leaf and flowering top of *Mentha spicata* Linné (*Mentha viridis* Linné) (Common Spearmint) or of *Mentha cardiaca* Gerard ex Baker (Scotch Spearmint) (Fam *Labiatae*).

Fresh spearmint is used in preparing mint sauce and also the well-known mint julep. The volatile oil is the only constituent of importance in this plant; the yield is from ½ to 1%.

Uses—A flavoring agent.

Spearmint Oil is the volatile oil distilled with steam from the fresh over-ground parts of the flowering plant of *Mentha spicata* or of *Mentha cardiaca;* contains not less than 55%, by volume, of $C_{10}H_{14}O$ (carvone = 150.22). The chief odoriferous constituent is the ketone *l-carvone*. American oil also contains *dihydracarveol acetate* [$CH_3COOC_{10}H_{17}$], *l-limonene* [$C_{10}H_{16}$], a small amount of *phellandrene* [$C_{10}H_{16}$] and traces of *esters of valeric* and *caproic acids*. Colorless, yellow or greenish yellow liquid, having the characteristic odor and taste of spearmint; specific gravity 0.917 to 0.934; soluble in 1 volume of 80% alcohol, but upon further dilution may become turbid. *Uses:* Primarily as a flavoring agent. It also has been used as a *carminative* in doses of 0.1 mL.

Sucrose

α-D-Glucopyranoside, β-D-fructofuranosyl-, Sugar; Cane Sugar; Beet Sugar

Sucrose [57-50-1] $C_{12}H_{22}O_{11}$ (342.30); a sugar obtained from *Saccharum officinarum* Linné (Fam *Graminae*), *Beta vulgaris* Linné (Fam *Chenopodiaceae*), and other sources. It contains no added substances.

For the structural formula, see page 382.

Preparation—Commercially from the sugar cane, beet root and sorghum. Originally, sugar cane was the only source, but at present the root of *Beta vulgaris* is used largely in Europe, and to an increasing degree in this country, for making sucrose.

The sugar cane is crushed and the juice amounting to about 80% is expressed with roller mills. The juice after "defecation" with lime and removal of excess of lime by carbonic acid gas, is run into vacuum pans for concentration and the saccharine juice is evaporated in this until it begins to crystallize. After the crystallization is complete, the warm mixture of crystals and syrup is run into centrifuges, in which the crystals of raw sugar are drained and dried. The syrup resulting as a by-product from raw sugar is known as *molasses*. Raw beet sugar is made by a similar process, but is more troublesome to purify than that made from sugar cane.

The refined sugar from either raw cane or beet sugar is prepared by dissolving the raw sugar in water, clarifying, filtering and, finally, decolorizing the solution by passing it through bone-black filters. The water-white solution finally is evaporated under reduced pressure to the crystallizing point and then forced to crystallize in small granules which are collected and drained in a centrifuge.

Description—Colorless or white crystals, crystalline masses or blocks, or a white, crystalline powder; odorless; sweet taste; stable in air; solutions neutral to litmus; melts with decomposition from 160 to 185°; specific gravity of about 1.57; specific rotation at 20° not less than +65.9°; unlike the other official sugars (dextrose, fructose and lactose), it does not reduce Fehling's solution even in hot solutions; also differs from these sugars in that it is darkened and charred by sulfuric acid in the cold; fermentable and, in dilute aqueous solutions, it ferments into alcohol and eventually acetic acid.

Sucrose is hydrolyzed by dilute mineral acids, slowly in the cold, and rapidly on heating into one molecule each of dextrose or levulose. This process is known technically as "inversion" and the product is referred to as "invert sugar;" the term inversion being derived from the change, through the hydrolysis, in the optical rotation from dextro of the sucrose to levo of the hydrolyzed product. The enzyme *invertase* also hydrolyzes sucrose.

Solubility—1 g in 0.5 mL of water, 170 mL of alcohol or in slightly more than 0.2 mL of boiling water; insoluble in chloroform or ether.

Uses—Principally as a pharmaceutical necessity for making syrups and lozenges. It gives viscosity and consistency to fluids.

Intravenous administration of hypertonic solutions has been employed chiefly to initiate *osmotic diuresis*. Such a procedure is not completely safe and renal tubular damage may result, particularly in patients with existing renal pathology. Safer and more effective diuretics are available.

Compressible Sugar

Sucrose that may contain some starch, malto-dextrin or invert sugar; contains 95.0 to 98.0% of sucrose.

Description—White, crystalline, odorless powder; sweet taste; stable in air.

Solubility—The sucrose portion is very soluble in water.

Uses—A *pharmaceutic aid* as a *tableting excipient* and *sweetening agent*. See also *Sucrose*.

Confectioner's Sugar

Sucrose ground together with corn starch to a fine powder; contains 95.0 to 97.0% of sucrose.

Description—Fine, white, odorless powder; sweet taste; stable in air; specific rotation not less than +62.6°.

Solubility—The sucrose portion is soluble in cold water; this is entirely soluble in boiling water.

Uses—A *pharmaceutic aid* as a *tableting excipient* and *sweetening agent*. See also *Sucrose*.

Syrup—page 1302.

Tolu Balsam

Tolu

A balsam obtained from *Myroxylon balsamum* (Linné) Harms (Fam *Leguminosae*).

Constituents—Up to 80% *resin*, about 7% *volatile oil*, 12 to 15% free *cinnamic acid*, 2 to 8% benzoic acid and 0.05% *vanillin*. The volatile oil is composed chiefly of *benzyl benzoate* and *benzyl cinnamate*, *ethyl benzoate*, *ethyl cinnamate*, a terpene called *tolene* (possibly identical with *phellandrene*) and the sesquiterpene alcohol *farnesol* also have been reported to be present.

Description—Brown or yellowish brown, plastic solid; transparent in thin layers and brittle when old, dried or exposed to cold temperatures; pleasant, aromatic odor resembling that of vanilla and a mild, aromatic taste.

Solubility—Nearly insoluble in water or in solvent hexane; soluble in alcohol, chloroform or ether, sometimes with slight residue or turbidity.

Uses—A *vehicle*, *flavoring agent* and stimulating *expectorant* as a syrup. It is also an ingredient of *Compound Benzoin Tincture* (page 760).

Tolu Balsam Syrup [Syrup of Tolu; Tolu Syrup]—*Preparation:* Add tolu balsam tincture (50 mL, all at once) to magnesium carbonate (10 g) and sucrose (60 g) in a mortar, and mix intimately. Gradually add purified water (430 mL) with trituration, and filter. Dissolve the remainder of sucrose (760 g) in the clear filtrate with gentle heating, strain the syrup while warm and add purified water (qs) through the strainer to make the product measure 1000 mL. Mix thoroughly. *Note:* May be made also in the following manner: Place the remaining sucrose (760 g) in a suitable percolator, the nock of which nearly is filled with loosely packed cotton, moistened after packing with a few drops of water. Pour the filtrate, obtained as directed in the formula above, upon the sucrose, and regulate the outflow to a steady drip of percolate. When all of the liquid has run through, return portions of the percolate, if necessary, to dissolve all of the sucrose. Then pass enough purified water through the cotton to make the product measure 1000 mL. Mix thoroughly. *Alcohol Content:* 3 to 5%. *Uses:* Chiefly for its agreeable *flavor* in cough syrups. *Dose:* 10 mL.

Tolu Balsam Tincture [Tolu Tincture]—*Preparation:* With tolu balsam (200 g), prepare a tincture by Process M (page 1543), using alcohol as the menstruum. *Alcohol Content:* 77 to 83%. *Uses:* A balsamic preparation employed as an addition to expectorant mixtures; also used in the preparation of *Tolu Balsam Syrup*. *Dose:* 2 mL.

Vanilla

Vanilla Bean

The cured, full-grown, unripe fruit of *Vanilla planifolia* Andrews, often known in commerce as Mexican or Bourbon Vanilla, or of *Vanilla tahitensis* J W Moore, known in commerce as Tahiti Vanilla (Fam *Orchidaceae*); yields not less than 12% of anhydrous extractive soluble in diluted alcohol.

Constituents—Contains a trace of a volatile oil, fixed oil, 4% resin, sugar, *vanillic acid* and about 2.5% *vanillin* (see below). The highest grade of vanilla comes from Madagascar; considerable quantities of the drug also are produced in Mexico.

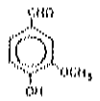
Uses—A flavor.

Note—Do not use if it has become brittle.

Vanilla Tincture [Extract of Vanilla]—*Preparation:* Add water (200 mL) to comminuted vanilla (cut into small pieces, 100 g) in a suitable covered container, and macerate during 12 hr, preferably in a warm place. Add alcohol (200 mL) to the mixture of vanilla and water, mix well and macerate about 3 days. Transfer the mixture to a percolator containing sucrose (in coarse granules, 200 g), and drain; then pack the drug firmly, and percolate slowly, using diluted alcohol (qs) as the menstruum. If the percolator is packed with an evenly distributed mixture of the comminuted vanilla, sucrose and clean, dry sand, the increased surface area permits more efficient percolation. This tincture is unusual in that it is the only official one in which sucrose is specified as an ingredient. *Alcohol Content:* 38 to 42%. *Uses:* A *flavoring agent*. See *Flavors*, page 1290.

Vanillin

Benzaldehyde, 4-hydroxy-3-methoxy-.



4-Hydroxy-3-methoxybenzaldehyde [121-33-5] $C_8H_8O_3$ (152.15).

Preparation—From vanilla, which contains 2 to 3%. It also is found in many other substances, including tissues of certain plants, crude beet sugar, asparagus and even asafetida. Commercially, it is made synthetically. While chemically identical with the product obtained from the "vanilla bean," "flavoring preparations" made from it never equal in flavor the preparation in which vanilla alone is used because vanilla contains other odorous products. It is synthesized by oxidation processes from either coniferin or eugenol, by treating guaiacol with chloroform in the presence of an alkali, and by other methods.

Description—Fine, white to slightly yellow crystals, usually needle-like having an odor and taste suggestive of vanilla; affected by light; solutions are acid to litmus; melts from 81 to 83°.

Solubility—1 g in about 100 mL of water, about 20 mL of glycerin or 20 mL of water at 80°; freely soluble in alcohol, chloroform, ether or solutions of the fixed alkali hydroxides.

Incompatibilities—Combines with *glycerin*, forming a compound which is almost insoluble in alcohol. It is decomposed by *alkalies* and is oxidized slowly by the *air*.

Uses—Only as a *flavor*. Solutions of it sometimes are sold as a synthetic substitute for vanilla for flavoring foods but it is inferior in flavor to the real vanilla extract.

Water—page 1300.

Water, Purified—page 1301.

Wild Cherry Syrup—page 1302.

Other Flavoring Agents

Anise NF IX [Anise Seed; European Aniseed; Sweet Cumin]—The dried ripe fruit of *Pimpinella anisum* Linné. It contains about 1.75% of volatile oil. *Uses:* A flavor and carminative.

Ceylon Cinnamon—The dried inner bark of the shoots of coppiced trees of *Cinnamomum zeylanicum* Nees (Fam *Lauraceae*); contains, in each 100 g, not less than 0.5 mL volatile oil. *Uses:* A *carminative* and *flavor*.

Clove—The dried flower-bud of *Eugenia caryophyllus* (Sprengel) Bullock et Harrison (Fam *Myrtaceae*). It contains, in each 100 g, not less than 16 mL of clove oil. *Uses:* An *aromatic* in doses of 0.25 g and as a condiment in foods.

Coriander—The dried ripe fruit of *Coriandrum sativum* Linné (Fam *Umbelliferae*); yields not less than 0.25 mL volatile coriander oil/100 g. *Uses:* Seldom used alone, but sometimes is combined with other agents, chiefly as a *flavor*. It also is used as a condiment and flavor in cooking.

Eucalyptol [Cineol; Cajeputol; $C_{10}H_{18}O$ (154.25)]—Obtained from eucalyptus oil and from other sources. Colorless liquid, having a characteristic, aromatic, distinctly camphoraceous odor and a pungent, cooling, spicy taste. 1 volume is soluble in 5 volumes of 60% alcohol; miscible with alcohol, chloroform, ether, glacial acetic acid or fixed or volatile oils; insoluble in water. *Uses:* Primarily as a *flavoring agent*. Locally it is employed for its *antiseptic* effect in inflammations of the nose and throat and in certain skin diseases. It sometimes is used by inhalation in bronchitis.

Fennel [Fennel Seed]—The dried ripe fruit of cultivated varieties of *Foeniculum vulgare* Miller (Fam *Umbelliferae*); contains 4 to 6% of an oxygenated volatile oil and 10% of a fixed oil. *Uses:* A *flavor* and *carminative*.

Ginger NF [Zingiber]—The dried rhizome of *Zingiber officinale* Roscoe (Fam *Zingiberaceae*), known in commerce as Jamaica Ginger, African Ginger and Cochin Ginger. The outer cortical layers often are removed either partially or completely. *Constituents:* A pungent substance, *gingerol*; volatile oil (Jamaica Ginger, about 1%; African Ginger, 2 to 3%), containing the terpenes *d-camphene* and *β-phellandrene* and the sesquiterpene *zingiberene*; *citral cineol* and *borneol*. *Uses:* A *flavoring agent*. It formerly was employed in a dose of 600 mg as an intestinal stimulant and carminative in colic and in diarrhea.

Ginger Oleoresin—Yields 18 to 35 mL of volatile ginger oil/100 g of oleoresin. *Preparation:* Extract the oleoresin from ginger, in moderately coarse powder, by percolation, using either acetone, alcohol or ether as the menstruum.

Glycyrrhiza Extract [Licorice Root Extract; Licorice]—An extract prepared from the rhizome and roots of species of *Glycyrrhiza* Tournefort ex Linné (Fam *Leguminosae*). *Description:* Brown powder or in flattened, cylindrical rolls or in masses; the rolls or masses have a glossy

black color externally, and a brittle, sharp, smooth, conchoidal fracture; the extract has a characteristic and sweet taste which is not more than very slightly acrid. *Uses:* A *flavoring agent*.

Lavendor [Lavandula]—The flowers of *Lavandula spica* (*Lavandula officinalis* or *Lavandula vera*); contains a volatile oil with the principal constituent *l*-linalyl acetate. *Uses:* A *perfume*.

Lemon Peel USP XV, BP [Fresh Lemon Peel]—The outer yellow rind of the fresh ripe fruit of *Citrus limon* (Linné) Burmann filius (Fam *Rutaceae*); contains a volatile oil and hesperidin. *Uses:* A *flavor*.

Lemon Tincture USP XVIII [Lemon Peel Tincture]—*Preparation:* From lemon peel, which is the outer yellow rind of the fresh, ripe fruit of *Citrus limon* (Linné) Burmann filius (Fam *Rutaceae*), by *Process M* (page 1543), 500 g of the peel being macerated in 900 mL alcohol and the preparation being completed with alcohol to make the product measure 1000 mL. Use talc as the filtering medium. The white portion of the rind must not be used, as the proportion of oil, which is found only in the yellow rind, is reduced and the bitter principle, hesperidin, introduced. *Alcohol Content:* 62 to 72%. *Uses:* A flavor, its fineness of flavor being assured as it comes from the fresh fruit, and being an alcoholic solution it is more stable than the oil.

Myrcia Oil [Bay Oil; Oil of Bay]—The volatile oil distilled from leaves of *Pimenta racemosa* (Miller) J W Moore (Fam *Myrtaceae*); contains the phenolic compounds eugenol and chavicol. *Uses:* In the preparation of bay rum as a *perfume*.

Orange Oil, Bitter—The volatile oil obtained by expression from the fresh peel of the fruit of *Citrus aurantium* Linné (Fam *Rutaceae*); contains primarily *d*-limonene. Pale yellow liquid with a characteristic, aromatic odor of the Seville orange; if it has a terebinthinate odor, it should not be dispensed; refractive index 1.4725 to 1.4755 at 20°. It differs little from *Orange Oil* (page 1296) except for the botanical source. Miscible with anhydrous alcohol and with about 4 volumes alcohol. *Uses:* A *flavor*.

Orange Peel, Bitter [Bitter Orange; Curacao Orange Peel; Bigarade Orange]—The dried rind of the unripe but fully grown fruit of *Citrus aurantium* Linné (Fam *Rutaceae*). *Constituents:* The inner part of the peel from the bitter orange contains a volatile oil and the glycoside *hesperidin* ($C_{28}H_{34}O_{15}$). This, upon hydrolysis in the presence of H_2SO_4, yields *hesperetin* ($C_{16}H_{14}O_6$), *rhamnose* ($C_6H_{12}O_5$), and D-glucose ($C_6H_{12}O_6$). *Uses:* A *flavoring agent*. It has been used as a bitter.

Orange Peel, Sweet USP XV—The fresh, outer rind of the non-artificially colored, ripe fruit of *Citrus sinensis* (Linné) Osbeck (Fam *Rutaceae*); the white, inner portion of the rind is to be excluded. Contains a volatile oil but no hesperidin, since the glycoside occurs in the white portion of the rind. *Uses:* A *flavor*.

Orris [Orris Root; Iris; Florentine Orris]—The peeled and dried rhizome of *Iris germanica* Linné, including its variety *florentina* Dykes (*Iris florentina* Linné), or of *Iris pallida* Lamarck (Fam *Iridaceae*); contains about 0.1 to 0.2% of a volatile oil (orris butter), myristic acid and the ketone irone; irone provides the fragrant odor of orris. *Uses:* A *perfume*.

Pimenta Oil [Pimento Oil; Allspice Oil]—The volatile oil distilled from the fruit of *Pimenta officinalis* Lindley (Fam *Myrtaceae*). *Uses:* A *carminative* and *stimulant* and also as a *condiment* in foods.

Rosemary Oil—The volatile oil distilled with steam from the fresh flowering tops of *Rosmarinus officinalis* Linné (Fam *Labiatae*); yields not less than 1.5% of esters calculated as bornyl acetate ($C_{12}H_{20}O_2$), and not less than 8% of total borneol ($C_{10}H_{18}O$), free and as esters. *Constituents:* The amount of esters, calculated as bornyl acetate, and of total borneol, respectively, varies somewhat with its geographic source. Cineol is present to the extent of about 19–25%, depending on the source. The terpenes *d*- and *l*-α-*pinene*, *dipentene* and *camphene*, and the ketone *camphor* also occur in this oil. *Description:* Colorless or pale yellow liquid, having the characteristic odor of rosemary, and a warm, camphoraceous taste; specific gravity 0.894 to 0.912. Soluble in 1 volume of 90% alcohol, by volume, but upon further dilution may become turbid. *Uses:* A *flavor* and *perfume*, chiefly, in rubefacient liniments such as *Camphor and Soap Liniment*.

Sassafras—The dried bark of the root of *Sassafras albidum* (Nuttall) Nees (Fam *Lauraceae*). *Uses:* Principally because of its high content of volatile oil which serves to disguise the taste of disagreeable substances. An infusion (*sassafras tea*) formerly was used extensively as a home remedy, particularly in the southern states.

Sassafras Oil—The volatile oil distilled with steam from *Sassafras*. *Uses:* A *flavor* by confectioners, particularly in hard candies. Either the oil or safrol is used as a *preservative* in mucilage and library paste, being far superior to methyl salicylate for this purpose. Since the oil is *antiseptic*, it sometimes is employed in conjunction with other agents for local application in diseases of the nose and throat; safrol also is used in this way.

Wild Cherry [Wild Black Cherry Bark]—The carefully dried stem bark of *Prunus serotina* Ehrhart (Fam *Rosaceae*), free of borke and preferably having been collected in autumn. *Constituents:* A glucoside of *d*-*mandelonitrile* ($C_6H_5CHOH.CN$) known as *prunasin* (page 385), the enzyme *emulsin*, tannin, a bitter principle, starch, resin, etc. In the BP and the English literature this drug has been termed "Virginian Prune"—a literal but incorrect translation of the older botanical name, *Prunus virginiana*. *Uses:* A *flavoring agent*, especially in cough preparations. It is an ingredient in *Wild Cherry Syrup*. As with bitter almond, contact with water, in the presence of emulsin, results in the production of benzaldehyde and HCN. All preparations of wild cherry should be made without heat in order to avoid destruction of the enzyme which is responsible for the production of the free active principles.

Diluting Agents

Diluting agents (vehicles or carriers) are indifferent substances which are used as solvents for active medicinals. They are of primary importance for diluting and flavoring drugs which are intended for oral administration, but a few such agents are designed specifically for diluting parenteral injections. The latter group is considered separately.

The expert selection of diluting agents has been an important factor in popularizing the "specialties" of manufacturing pharmacists. Since a large selection of diluting agents is available in a choice of colors and flavors, the prescriber has an opportunity to make his own prescriptions more acceptable to the patient. The best diluting agent is usually the best solvent for the drug. Water-soluble substances, for example, should be flavored and diluted with an aqueous agent and alcohol-soluble drugs with an alcoholic vehicle. Thus, the diluting agents presented herein are divided into three groups on the basis of their physical properties: aqueous, hydroalcoholic and alcoholic.

Aqueous Diluting Agents

Aqueous diluting agents include aromatic waters, syrups and mucilages. Aromatic waters are used as diluting agents for water-soluble substances and salts, but cannot mask the taste of very disagreeable drugs. Some of the more common flavored aqueous agents and the official forms of water are listed below.

Orange Flower Water

Stronger Orange Flower Water; Triple Orange Flower Water

A saturated solution of the odoriferous principles of the flowers of *Citrus aurantium* Linné (Fam. *Rutaceae*), prepared by distilling the fresh flowers with water and separating the excess volatile oil from the clear, water portion of the distillate.

Description—Should be nearly colorless, clear or only faintly opalescent; the odor should be that of the orange blossoms; it must be free from empyreuma, mustiness and fungoid growths.

Uses—A *vehicle flavor* and *perfume* in syrups, elixirs and solutions.

Peppermint Water

A clear, saturated solution of peppermint oil in purified water, prepared by one of the processes described under *Aromatic Waters* (page 1522).

Uses—A *carminative* and *flavored vehicle*.

Dose—15 mL.

Tolu Balsam Syrup—page 1209.

Water

Water [7732-18-5] H_2O (18.02).

Drinking water, which is subject to EPA regulations with respect to drinking water, and which is delivered by the municipal or other local public system or drawn from a private well or reservoir, is the starting material for all forms of water covered by Pharmacopeial monographs.

Drinking water may be used in the preparation of USP drug substances (eg, in the extraction of certain vegetable drugs and in the manufacture of a few preparations used externally) but not in the preparation of dosage forms, or in the preparation of reagents or test solutions. It is no longer the subject of a separate monograph (in the USP), inasmuch as the cited standards vary from one community to another and generally are beyond the control of private parties or corporations.

Purified Water

Water obtained by distillation, ion-exchange treatment, reverse osmosis or any other suitable process; contains no added substances.

Caution—Do not use this in preparations intended for parenteral administration. For such purposes, use Water for Injection, Bacteriostatic Water for Injection, or *Sterile Water for Injection,* page 1304.

Preparation—From water complying with EPA regulations with respect to drinking water. A former official process for water, when prepared by distillation, is given below. The pharmacist who is preparing sterile solutions, and must have freshly distilled water of exceptionally high grade, not only free from all bacterial or other microscopic growths but also free from the products of metabolic processes resulting from the growth of such organisms in the water, advantageously may follow this plan. The metabolic products commonly are spoken of as pyrogens and usually consist of complex organic compounds which cause febrile reactions if present in the solvent for parenteral medicinal substances.

Distillation Process

Water	1000 Vol
To make	750 Vol

Distil the water from a suitable apparatus provided with a block-tin or glass condenser. Collect the first 100 volumes and reject this portion. Then collect 750 volumes and keep the distilled water in glass-stoppered bottles, which have been rinsed with steam or very hot distilled water immediately before being filled. The first 100 volumes are discarded to eliminate foreign volatile substances found in ordinary water and only 750 volumes are collected, since the residue in the still contains concentrated dissolved solids.

Description—Colorless, clear liquid, without odor or taste.

Uses—A *pharmaceutic aid* (vehicle and solvent). It must be used in compounding dosage forms for internal (oral) administration as well as sterile pharmaceuticals applied externally, such as collyria and dermatological preparations, but these must be sterilized before use.

Whenever water is called for in official tests and assays, this must be used.

Syrups Used as Diluting Agents

Syrups are useful as diluting agents for water-soluble drugs and act both as solvents and flavoring agents. The flavored syrups usually consist of simple syrup (85% sucrose in water) containing appropriate flavoring substances. *Glycyrrhiza Syrup* is an excellent vehicle for saline substances because of its colloidal properties, sweet flavor and lingering taste of licorice. *Acacia Syrup* is valuable in disguising the taste of urea. Fruit syrups are especially effective for masking sour tastes. *Aromatic Eriodictyon Syrup* is the diluting agent of choice for masking the bitter taste of alkaloids. *Cocoa Syrup* and *Cherry Syrup* are good general flavoring agents.

Acacia Syrup

Acacia, granular or powdered	100 g
Sodium Benzoate	1 g
Vanilla Tincture	5 mL
Sucrose	800 g
Purified Water, a sufficient quantity,	
To make	1000 mL

Mix the acacia, sodium benzoate and sucrose; then add 425 mL of purified water, and mix well. Heat the mixture on a steam bath until solution is completed. When cool, remove the scum, add the vanilla tincture and sufficient purified water to make the product measure 1000 mL, and strain, if necessary.

Uses—A *flavored vehicle* and *demulcent*.

Cherry Syrup

Syrupus Cerasi

Cherry Juice	475 mL
Sucrose	800 g
Alcohol	20 mL
Purified Water, a sufficient quantity,	
To make	1000 mL

Dissolve the sucrose in cherry juice by heating on a steam bath, cool and remove the foam and floating solids. Add the alcohol and sufficient purified water to make 1000 mL, and mix.

Alcohol Content—1 to 2%.

Uses—A pleasantly *flavored vehicle* which is particularly useful in masking the taste of saline and sour drugs.

Cocoa Syrup

Cacao Syrup; Chocolate-flavored Syrup; Chocolate Syrup

Cocoa	180	g
Sucrose	600	g
Liquid Glucose	180	g
Glycerin	50	mL
Sodium Chloride	2	g
Vanillin	0.2	g
Sodium Benzoate	1	g
Purified Water, a sufficient quantity,		
To make	1000	mL

Mix the sucrose and the cocoa, and to this mixture gradually add a solution of the liquid glucose, glycerin, sodium chloride, vanillin and sodium benzoate in 325 mL of hot purified water. Bring the entire mixture to a boil, and maintain at boiling temperature for 3 min. Allow to cool to room temperature and add sufficient purified water to make the product measure 1000 mL.

Note—Cocoa containing not more than 12% nonvolatile, ether-soluble extractive ("fat") yields a syrup having a minimum tendency to separate. "Breakfast cocoa" contains over 22% "fat."

Uses—A pleasantly *flavored vehicle*.

Aromatic Eriodictyon Syrup

Aromatic Yerba Santa Syrup; Syrupus Corrigens

Eriodictyon Fluidextract	32	mL
Potassium Hydroxide Solution (1 in 30)	25	mL
Compound Cardamom Tincture	65	mL
Lemon Oil	0.5	mL
Clove Oil	1	mL
Alcohol	32	mL
Sucrose	800	g
Magnesium Carbonate	5	g
Purified Water, a sufficient quantity,		
To make	1000	mL

Dissolve the oils in the alcohol, add the fluidextract and the tincture, then the potassium hydroxide solution and 325 mL of purified water. Add the magnesium carbonate, shake the mixture, allow it to stand overnight, filter and add sufficient purified water through the filter to make the liquid measure 500 mL. Pour this filtrate upon the sucrose contained in a bottle, dissolve by placing the bottle in hot water and agitating the contents frequently. Cool the solution and add sufficient purified water to make the product measure 1000 mL.

Alcohol Content—6 to 8%.

Incompatibilities—Alkaline in reaction due to the potassium hydroxide used in its manufacture. *Acids* are neutralized with usually a

concurrent precipitation of the resins of the syrup. The tannin which it contains introduces the incompatibilities of that substance.

Uses—A pleasantly *flavored vehicle*, especially adapted to the administration of bitter substances like quinine.

Syrup

Simple Syrup

Sucrose	850 g
Purified Water, a sufficient quantity,	
To make	1000 mL

May be prepared by using boiling water or, preferably, without heat, by the following process:

Place the sucrose in a suitable percolator the neck of which is nearly filled with loosely packed cotton moistened, after packing, with a few drops of water. Pour carefully about 460 mL of purified water upon the sucrose, and regulate the outflow to a steady drip of percolate. Return the percolate, if necessary, until all of the sucrose has dissolved. Then wash the inside of the percolator and the cotton with sufficient purified water to bring the volume of the percolate to 1000 mL, and mix.

Specific Gravity—Not less than 1.30.

Uses—A *sweet vehicle*, sweetening agent and as the basis for many flavored and medicated syrups.

Other Syrups Used As Diluting Agents

Citric Acid Syrup USP XVIII [Syrup of Lemon]—*Preparation:* Dissolve citric acid (hydrous, 10 g) in purified water (10 mL), and mix the solution with syrup (950 mL). Add lemon tincture (10 mL), and enough syrup to make the product measure 1000 mL, and mix. *Note: Do not dispense it if it has a terebinthine odor or taste or shows other indications of deterioration. Alcohol Content:* Less than 1%. *Incompatibilities:* Reactions characteristic of the acid which it contains; hence, it is not a suitable vehicle for alkaline ingredients such as phenobarbital sodium from which it precipitates phenobarbital. *Uses:* Solely as a *pleasant vehicle*, the formula making it possible to prepare extemporaneously and quickly a syrup having the flavor of lemon.

Glycyrrhiza Syrup USP XVIII [Licorice Syrup]—*Preparation:* Add fennel oil (0.05 mL) and anise oil (0.5 mL) to glycyrrhiza fluidextract (250 mL) and agitate until mixed. Then add syrup (qs) to make the product measure 1000 mL, and mix. *Alcohol Content:* 5 to 6%. *Incompatibilities:* The characteristic flavor is destroyed by acids due to a precipitation of the glycyrrhizin. *Uses:* A *flavored vehicle*, especially adapted to the administration of bitter or nauseous substances.

Hydriodic Acid Syrup—Contains, in each 100 mL 1.3 to 1.5 g HI (127.91). *Preparation:* Mix diluted hydriodic acid (140 mL) with purified water (550 mL), and dissolve dextrose (450 g) in this mixture by agitation. Add purified water (qs) to make the product measure 1000 mL, and filter. *Caution: It must not be dispensed if it contains free iodine, as evidenced by a red coloration. Description:* Transparent, colorless, or not more than pale straw-colored, syrupy liquid; odorless and has a sweet, acidulous taste; specific gravity about 1.18; hydriodic acid is decomposed easily in simple aqueous solution (unless protected by hypophosphorous acid) free iodine being liberated, and if taken internally, when in this condition, it is irritating to the alimentary tract. The dextrose used in this syrup should be of the highest grade obtainable. *Incompatibilities:* The reactions of the *acids* (page 1523) as well as those of the water-soluble iodide salts. Oxidizing agents liberate iodine; alkaloids may be precipitated. *Uses:* Traditionally as a *vehicle for expectorant* drugs. Its therapeutic properties are those of the iodides. *Dose: Usual*, 5 mL.

Raspberry Syrup USP XVIII—*Preparation:* Dissolve sucrose (800 g) in raspberry juice (475 mL) by heating on a steam bath, cool and remove the foam and floating solids. Add alcohol (20 mL) and purified water (qs) to make 1000 mL, and mix. *Alcohol Content:* 1 to 2%. *Incompatibilities:* Raspberry juice is prepared to contain not less than 1.5% citric acid; the syrup, therefore, has reactions characteristic of this acid, notably its incompatibility with alkaline substances. *Uses:* A pleasantly *flavored vehicle* used to disguise the salty or sour taste of saline medicaments.

Wild Cherry Syrup USP XVIII—*Preparation:* Pack wild cherry (in coarse powder, 150 g), previously moistened with water (100 mL), in a cylindrical percolator, and add water (qs) to leave a layer of it above the powder. Macerate for 1 hr, then proceed with rapid percolation, using added water, until 400 mL of percolate is collected. Filter the percolate, if necessary, add sucrose (675 g) and dissolve it by agitation, then add glycerin (150 mL), alcohol (20 mL) and water (qs) to make the product measure 1000 mL. Strain if necessary. *It may be made also in the following manner:* The sucrose may be dissolved by placing it in a second percolator as directed for preparing *Syrup*, and allowing the percolate from the wild cherry to flow through it and into a graduated vessel containing the glycerin and alcohol until the total volume measures 1000 mL. *Note:* Heat is avoided, lest the enzyme emulsin be inactivated. If this should happen, the preparation would contain no free HCN, upon which its action as a sedative for coughs mainly depends. For a discussion of the chemistry involved, see *Wild Cherry* (page 1300). *Alcohol Content:* 1 to 2%. *Uses:* Chiefly as a *flavored vehicle* for cough syrups.

Mucilages Used as Diluting Agents

Mucilages are also suitable as diluting agents for water-soluble substances, and are especially useful in stabilizing suspensions and emulsions.

The following mucilage used for this purpose is described under *Emulsifying and Suspending Agents*, page 1304.

Acacia Mucilage—page 1304.

Hydroalcoholic Diluting Agents

Hydroalcoholic diluting agents are suitable for drugs soluble in either water or diluted alcohol. The most important agents in this group are the elixirs. These solutions contain approximately 25% alcohol. *Medicated* elixirs which have therapeutic activity in their own right are not included in this section. Listed below are the common, nonmedicated elixirs which are used purely as diluting agents or solvents for drugs.

Aromatic Elixir

Simple Elixir

Orange Oil	2.4 mL
Lemon Oil	0.6 mL
Coriander Oil	0.24 mL
Anise Oil	0.06 mL
Syrup	375 mL
Talc	30 g
Alcohol,	
Purified Water, each, a sufficient quantity,	
To make	1000 mL

Dissolve the oils in alcohol to make 250 mL. To this solution add the syrup in several portions, agitating vigorously after each addition, and afterwards add, in the same manner, the required quantity of purified water. Mix the talc with the liquid, and filter through a filter wetted with diluted alcohol, returning the filtrate until a clear liquid is obtained.

Alcohol Content—21 to 23%.

Uses—A pleasantly *flavored vehicle*, employed in the preparation of many other elixirs. The chief objection to its extensive use is the high alcohol content (about 22%) which at times may counteract the effect of other medicines.

Cardamom Spirit, Compound—RPS-15, page 1236.

Other Hydroalcoholic Diluting Agents

Glycyrrhiza Elixir [Elixir Adjuvans; Licorice Elixir]—*Preparation:* Mix glycyrrhiza fluidextract (125 mL) and aromatic elixir (875 mL) and filter. *Alcohol Content:* 21 to 23%. *Uses:* A *flavored vehicle.*

Flavored Alcoholic Solutions

Flavored alcoholic solutions, of high alcoholic concentration, are useful as flavors to be added in small quantities to syrups or elixirs. The alcohol content of these solutions is approximately 50%. There are two types of flavored alcoholic solutions: tinctures and spirits. Only nonmedicated tinctures and spirits are used as flavoring agents.

Compound Cardamom Tincture

Cardamom Seed, in moderately coarse powder	20 g
Cinnamon, in fine powder	25 g
Caraway, in moderately coarse powder	12 g
To make	1000 mL

Prepare a tincture by Process M (page 1543), macerating the mixed powders in 750 mL of a mixture of 50 mL of glycerin and 950 mL of diluted alcohol and completing the preparation by using first the remainder of the mixture of alcohol and glycerin prepared as directed above, and then diluted alcohol.

Note—Compound cardamom tincture may be colored with one or more colors (page 1288).

Alcohol Content—43 to 47%.

Uses—A useful vehicle because of its pleasant *flavor* and *color*.

Lemon Tincture—page 1300.

Myrcia Spirit, Compound—RFS-13, page 452.

Orange Spirit, Compound—page 1296.

Orange Peel, Sweet, Tincture—page 1296.

Peppermint Spirit—page 798.

Diluting Agents for Injections

Injections are liquid preparations, usually solutions or suspensions of drugs, intended to be injected through the skin into the body. Diluting agents used for these preparations may be aqueous or nonaqueous and must meet the requirements for sterility and also of the pyrogen test. Aqueous diluting agents include such preparations as *Sterile Water for Injection* and various sterile, aqueous solutions of electrolytes and/or dextrose. Nonaqueous diluting agents are generally fatty oils of vegetable origin, fatty esters and polyols such as propylene glycol and polyethylene glycol. These agents are used to dissolve or dilute oil-soluble substances and to suspend water-soluble substances when it is desired to decrease the rate of absorption and, hence, prolong the duration of action of the drug substances. Preparations of this type are given intramuscularly. See *Parenteral Preparations*, page 1545.

Corn Oil

Maize Oil

The refined fixed oil obtained from the embryo of *Zea mays* Linné (Fam *Gramineae*).

Preparation—Expressed from the Indian corn embryos or germs separated from the grain in starch manufacture.

Description—Clear, light yellow, oily liquid with a faint characteristic odor and taste; specific gravity 0.914 to 0.921.

Solubility—Slightly soluble in alcohol; miscible with ether, chloroform, benzene or solvent hexane.

Uses—Main official use is as a *solvent* and *vehicle for injections*. It is used as an edible oil substitute for solid fats in the management of hypercholesterolemia. Other uses include making soaps and for burning. It is a semidrying oil and therefore unsuitable for lubricating or mixing paint.

Cottonseed Oil

Cotton Seed Oil; Cotton Oil

The refined fixed oil obtained from the seed of cultivated plants of various varieties of *Gossypium hirsutum* Linné or of other species of *Gossypium* (Fam *Malvaceae*).

Preparation—Cotton seeds contain about 15% oil. The testae of the seeds are first separated, and the kernels are subjected to high pressure in hydraulic presses. The crude oil thus has a bright red to blackish red color. It requires purification before it is suitable for medicinal or food purposes.

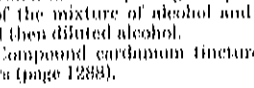
Description—Pale yellow, oily liquid with a bland taste; odorless or nearly so; particles of solid fat may separate below 10°; solidifies at about 0° to −5°; specific gravity 0.915 to 0.921.

Solubility—Slightly soluble in alcohol; miscible with ether, chloroform, solvent hexane or carbon disulfide.

Uses—Official as a *solvent* and *vehicle for injections*. It is sometimes taken orally as a mild cathartic in the dose of 30 mL or more. Taken internally, digestible oils retard gastric secretion and motility and increase the caloric intake. It also is used in the manufacture of soaps, oleomargarine, lard substitutes, glycerin, lubricants and cosmetics.

Ethyl Oleate

(*Z*)-9-Octadecenoic acid, ethyl ester



Ethyl oleate [111-62-6] $C_{20}H_{38}O_2$ (310.52).

Preparation—Among other ways, by reacting ethanol with oleoyl chloride in the presence of a suitable dehydrochlorinating agent.

Description—Mobile, practically colorless liquid, having an agreeable taste; specific gravity 0.868 to 0.874; acid value not greater than 0.5; iodine value 75 to 95; sterilized by heating at 150° for 1 hr; properties similar to those of almond and arachis oils, but is less viscous and more rapidly absorbed by the tissues; boils about 207°.

Solubility—Does not dissolve in water; miscible with vegetable oils, mineral oil, alcohol or most organic solvents.

Uses—A *vehicle* for certain intramuscular injectable preparations.

Peanut Oil

Arachis Oil; Groundnut Oil; Nut Oil; Earth-Nut Oil

The refined fixed oil obtained from the seed kernels of one or more of the cultivated varieties of *Arachis hypogaea* Linné (Fam *Leguminosae*).

Description—Colorless or pale yellow, oily liquid, with a characteristic nutty odor and a bland taste; specific gravity 0.912 to 0.920.

Solubility—Very slightly soluble in alcohol; miscible with ether, chloroform or carbon disulfide.

Uses—A *solvent* in preparing oil solutions for injection (page 1549). It also is used for making liniments, ointments, plasters and soaps, as a substitute for olive oil.

Sesame Oil

Teel Oil; Benne Oil; Gingili Oil

The refined fixed oil obtained from the seed of one or more cultivated varieties of *Sesamum indicum* Linné (Fam *Pedaliaceae*).

Description—Pale yellow, almost odorless, oily liquid with a bland taste; specific gravity 0.916 to 0.921.

Solubility—Slightly soluble in alcohol; miscible with ether, chloroform, solvent hexane or carbon disulfide.

Uses—A *solvent* and *vehicle* in official injections. It is used much like olive oil both medicinally and for food. It does not readily turn rancid. It also is used in the manufacture of cosmetics, iodized oil, liniments, ointments and oleomargarine.

Water for Injection

Water purified by distillation or by reverse osmosis. It contains no added substance.

Caution—It is intended for use as a solvent for the preparation of parenteral solutions. For parenteral solutions that are prepared under aseptic conditions and are not sterilized by appropriate filtration or in the final container, first render it sterile and thereafter protect it from microbial contamination.

Description—Clear, colorless, odorless liquid.

Uses—*Pharmaceutic aid* (vehicle and solvent).

Bacteriostatic Water for Injection

Sterile water for injection containing one or more suitable antimicrobial agents.

Note—Use it with due regard for the compatibility of the antimicrobial agent or agents it contains with the particular medicinal substance that is to be dissolved or diluted.

Uses—*Sterile vehicle* for parenteral preparations.

Sterile Water for Injection

Water for Parenterals

Water for injection sterilized and suitably packaged. It contains no antimicrobial agent or other added substance.

Description—Clear, colorless, odorless, liquid.

Uses—For the preparation of *all aqueous parenteral solutions*, including those used in *animal assays*. See page 1547 for a detailed discussion.

Sterile Water for Irrigation

Water for injection that has been sterilized and suitably packaged. It contains no antimicrobial agent or other added substance.

Description—Clear, colorless, odorless liquid.

Uses—An *irrigating solution*.

Emulsifying and Suspending Agents

An emulsion is a two-phase system in which one liquid is dispersed in the form of small globules throughout another liquid that is immiscible with the first liquid. Emulsions are formed and stabilized with the help of emulsifying agents, which are surfactants and/or viscosity-producing agents. A suspension is defined as a preparation containing finely divided insoluble material suspended in a liquid medium. The presence of a suspending agent is required to overcome agglomeration of the dispersed particles and to increase the viscosity of the medium so that the particles settle more slowly. Emulsifying and suspending agents are used extensively in the formulation of elegant pharmaceutical preparations for oral, parenteral and external use. For the theoretical and practical aspects of emulsions the interested reader is referred to pages 300 and 1605. More detailed information on the use of suspending agents is given on page 1538.

Acacia

Gum Arabic

The dried gummy exudate from the stems and branches of *Acacia senegal* (Linné) Willdenow or of other related African species of *Acacia* (Fam *Leguminosae*).

Constituents—Principally calcium, magnesium and potassium salts of the polysaccharide *arabic acid*, which on acid hydrolysis yields L-arabinose, L-rhamnose, D-galactose and an aldobionic acid containing D-glucuronic acid and D-galactose.

Description—*Acacia:* Spheroidal tears up to 32 mm in diameter or angular fragments of white to yellowish white color; translucent or somewhat opaque; very brittle; almost odorless; produces a mucilaginous sensation on the tongue. *Flake Acacia:* White to yellowish white, thin flakes. *Powdered Acacia:* White to yellowish white, angular microscopic fragments. *Granular Acacia:* White to pale yellowish white, fine granules. *Spray-dried Acacia:* White to off-white compacted microscopic fragments or whole spheres.

Solubility—Insoluble in alcohol, but almost completely soluble in twice its weight of water at room temperature; the resulting solution flows readily and is acid to litmus.

Incompatibilities—*Alcohol or alcoholic solutions* precipitate acacia as a stringy mass when the alcohol amounts to more than about 35% of the total volume. Solution is effected by dilution with water. The mucilage is destroyed through precipitation of the acacia by *heavy metals*. *Borax* also causes a precipitation which is prevented by glycerin. It contains calcium and, therefore, possesses the incompatibilities of this ion.

It contains a *peroxidase* which acts as an oxidizing agent and produces colored derivatives of *aminopyrine, antipyrine, cresol, guaiacol, phenol, tannin, thymol, vanillin* and other substances. Among the alkaloids affected are *atropine, apomorphine, cocaine, homatropine, hyoscyamine, morphine, physostigmine* and *scopolamine*. A partial destruction of the alkaloid occurs in the reaction. Heating the solution of acacia for a few minutes at 100° destroys the peroxidase and the color reactions are avoided.

Uses—Extensively as a *suspending agent* for insoluble substances in water (page 1538), in the preparation of emulsions (pages 298 and 1534) and for making pills and troches (page 1664).

It is used for its *demulcent* action in inflammations of the throat or stomach.

Its solutions should not be used as a substitute for serum protein in the treatment of *shock* and as a *diuretic* in hypoproteinemic edema, since it produces serious syndromes that may result in death.

Acacia Mucilage [Mucilage of Gum Arabic]—*Preparation:* Place acacia (in small fragments, 350 g) in a graduated bottle having a wide mouth and a capacity not greatly exceeding 1000 mL, wash the drug with cold purified water, allow it to drain and add enough warm purified water, in which benzoic acid (2 g) has been dissolved, to make the product measure 1000 mL. After stoppering, lay the bottle on its side, rotate it occasionally, and when the acacia has dissolved strain the mucilage. *It also may be prepared as follows:* dissolve benzoic acid (2 g) in purified water (400 mL) with the aid of heat, and add the solution to powdered or granular acacia (350 g), in a mortar, triturating until the acacia is dissolved. Then add sufficient purified water to make the product measure 1000 mL, and strain if necessary. This second method is primarily for extemporaneous preparation. *Uses:* A *demulcent* and a *suspending agent*. It also has been employed as an *excipient* in making pills and troches, and as an *emulsifying agent* for cod liver oil and other substances. *Caution—It must be free from mold or any other indication of decomposition.*

Agar

Agar-Agar; Vegetable Gelatin; Gelose; Chinese or Japanese Gelatin

The dried, hydrophilic, colloidal substance extracted from *Gelidium cartilagineum* (Linné) Gaillon (Fam *Gelidiaceae*), *Gracilaria confervoides* (Linné) Greville (Fam *Sphaerococcaceae*) and related red algae (Class *Rhodophyceae*).

Constituents—Chiefly of the calcium salt of a galactan mono-(acid sulfate).

Description—Usually in bundles of thin, membranous, agglutinated strips or in cut, flaked, or granulated forms; may be weak yellowish orange, yellowish gray to pale yellow or colorless; tough when damp, brittle when dry; odorless or with a slight odor; produces a mucilaginous sensation on the tongue. Also supplied as a white to yellowish white or pale-yellow powder.

Solubility—Insoluble in cold water; soluble in boiling water.

Incompatibilities—Like other gums, it is dehydrated and precipitated from solution by *alcohol*. *Tannic acid* causes precipitation; *electrolytes* cause partial dehydration and decrease in viscosity of sols.

Uses—A relatively ineffective bulk-producing laxative used in a variety of proprietary cathartics. In mineral oil emulsions it acts as a stabilizer. The usual dose is 4 to 16 g once or twice a day.

It also is used in culture media for bacteriological work and in the manufacture of ice cream, confectionaries, etc.

Alginic Acid

Alginic acid [9005-32-7] (average equivalent weight 200); a hydrophilic colloidal carbohydrate extracted with dilute alkali from various species of brown seaweeds (*Phaeophyceae*).

Preparation—Precipitates when an aqueous solution of *Sodium Alginate* is treated with mineral acid.

Description—White to yellowish white, fibrous powder; odorless or practically odorless, and tasteless; pH (3 in 100 dispersion in water) 1.5 to 3.5; pK_a (0.1*N* NaCl, 20°) 3.42.

Solubility—Insoluble in water or organic solvents; soluble in alkaline solutions.

Uses—A *pharmaceutic aid* (tablet binder and emulsifying agent). It is used as a sizing agent in the paper and textile industries.

Sodium Alginate

Alginic acid, sodium salt; Algin; Manucol; Norgine; Kelgin (*Kelco*)

Sodium alginate [9005-38-3] (average equivalent weight 220); the purified carbohydrate product extracted from brown seaweeds by the use of dilute alkali. It consists chiefly of the sodium salt of alginic acid, a polyuronic acid composed of beta-D-mannuronic acid residues linked so that the carboxyl group of each unit is free while the aldehyde group is shielded by a glycosidic linkage.

Description—Nearly odorless and tasteless, coarse or fine powder, yellowish white in color.

Solubility—Dissolves in water, forming a viscous, colloidal solution; insoluble in alcohol or in hydroalcoholic solutions in which the alcohol content is greater than about 30% by weight; insoluble in chloroform, ether or acids, when the pH of the solution becomes lower than about 3.

Uses—A *thickening* and *emulsifying agent*. This property makes it useful in a variety of areas. For example, it is used to impart smoothness and body to ice cream and to prevent formation of ice particles.

Bentonite

Wilhinite; Soap Clay; Mineral Soap

Bentonite [1302-78-9]; a native, colloidal, hydrated aluminum silicate.

Occurrence—Bentonite is found in the Midwest of the US and Canada. Originally called *Taylorite* after its discoverer in Wyoming, its name was changed to bentonite after its discovery in the Fort Benton formation of the Upper Cretaceous of Wyoming.

Description—Very fine, odorless powder with a slightly earthy taste, free from grit; the powder is nearly white, but may be a pale buff or cream-colored.

The US Geological Survey has defined bentonite as "a transported stratified clay formed by the alteration of volcanic ash shortly after deposition." Chemically, it is $Al_2O_3.4SiO_2.H_2O$ plus other minerals as impurities. It consists of colloidal crystalline plates, of less than microscopic dimensions in thickness, and of colloidal dimensions in breadth. This fact accounts for the extreme swelling that occurs when it is placed in water, since the water penetrates between an infinite number of plates. A good specimen swells 12 to 14 times its volume.

Solubility—Insoluble in water or acids, but it has the property of absorbing large quantities of water, swelling to approximately 12 times its original volume, and forming highly viscous thixotropic suspensions or *gels*. This property makes it highly useful in pharmacy. Its gel-forming property is augmented by the addition of small amounts of alkaline substances, such as magnesium oxide. It does not swell in organic solvents.

Incompatibilities—*Acids* and *acid salts* decrease its water-absorbing power and thus cause a breakdown of the magma. Suspensions are most stable at a pH above 7.

Uses—A *protective colloid* for the *stabilization of suspensions*. It also has been used as an emulsifier for oil and as a base for plasters, ointments and similar preparations.

Bentonite Magma—*Preparation:* Sprinkle bentonite (50 g), in portions, on hot purified water (800 g), allowing each portion to become thoroughly wetted without stirring. Allow it to stand with occasional stirring for 24 hr. Stir until a uniform magma is obtained, add purified water to make 1000 g, and mix. The magma may be prepared also by mechanical means such as by use of a blender, as follows: Place purified water (about 500 g) in the blender, and while the machine is running, add bentonite (50 g). Add purified water to make up to about 1000 g or up to the operating capacity of the blender. Blend the mixture for 5 to 10 min, add purified water to make 1000 g, and mix. *Uses:* A *suspending agent* for insoluble medicaments.

Carbomer

Carboxypolymethylene

A synthetic high-molecular-weight cross-linked polymer of acrylic acid; contains 56 to 68% of carboxylic acid (—COOH) groups. The viscosity of a neutralized preparation (2.5 g/500 mL water) is 30,000 to 40,000 centipoises.

Description—White, fluffy powder with a slight characteristic odor; hygroscopic; pH (1 in 100 dispersion) about 3; specific gravity about 1.41.

Solubility (neutralized with alkali hydroxides or amines)—Dissolves in water, alcohol and glycerin.

Uses—A *thickening, suspending, dispersing* and *emulsifying agent* for pharmaceuticals, cosmetics, waxes, paints and other industrial products.

Carrageenan

Carrageenan [9000-07-1].

Preparation—The hydrocolloid extracted with water or aqueous alkali from certain red seaweeds of the class *Rhodophyceae*, and separated from the solution by precipitation with alcohol (methanol, ethanol or isopropanol) or by drum-roll drying or freezing.

Constituents—It is a variable mixture of potassium, sodium, calcium, magnesium and ammonium sulfate esters of galactose and 3,6-anhydrogalactose copolymers, the hexoses being alternately linked α-1,3 and β-1,4 in the polymer. The three main types of copolymers present are *kappa-carrageenan*, *iota-carrageenan* and *lambda-carrageenan*, which differ in the composition and manner of linkage of monomeric units and the degree of sulfation (the ester sulfate content for carrageenans varies from 18 to 40%). *Kappa*-carrageenan and *iota*-carrageenan are the gelling fractions; *lambda*-carrageenan is the nongelling fraction. The gelling fractions may be separated from the nongelling fraction by addition of potassium chloride to an aqueous solution of carrageenan. Carrageenan separated by drum-roll drying may contain mono- and diglycerides or up to 5% of polysorbate 80 used as roll-stripping agents.

Description—Yellow-brown to white, coarse to fine powder; odorless; tasteless, producing a mucilaginous sensation on the tongue.

Solubility—All carrageenans hydrate rapidly in cold water, but only *lambda*-carrageenan and sodium carrageenans dissolve completely. Gelling carrageenans require heating to about 80° for complete solution where potassium and calcium ions are present.

Uses—In the pharmaceutical and food industries as an emulsifying, suspending and gelling agent.

Carboxymethylcellulose Sodium

Carbose D; Carboxymethocel S; CMC; Cellulose Gum (*Hercules*)

Cellulose, carboxymethyl ether, sodium salt [9004-32-4]; contains 6.5–9.5% of sodium (Na), calculated on the dried basis. It is available in several viscosity types: low, medium, high and extra high.

Description—White to cream-colored powder or granules; the powder is hygroscopic; pH (1 in 100 aqueous solution) about 7.5.

Solubility—Easily dispersed in water to form colloidal solutions; insoluble in alcohol, ether or most other organic solvents.

Uses—*Pharmaceutic aid* (suspending agent, tablet excipient or viscosity-increasing agent). In tablet form it is used as a hydrophilic colloid laxative.

Dose—*Usual, adult, laxative*, 1.5 g 3 or 4 times a day.

Dosage Form: Tablets: 500 mg.

Powdered Cellulose

Cellulose [9004-34-6] $(C_6H_{10}O_5)_n$; purified, mechanically disintegrated cellulose prepared by processing alpha cellulose obtained as a pulp from fibrous plant materials.

Description—White, odorless substance, consisting of fibrous particles, which may be compressed into self-binding tablets which disintegrate rapidly in water; exists in various grades, exhibiting degrees of fineness ranging from a free-flowing dense powder to a coarse, fluffy, nonflowing material; pH (supernatant liquid of a 10 g/90 mL aqueous suspension after 1 hr) 5 to 7.5.

Solubility—Insoluble in water, dilute acids or nearly all organic solvents; slightly soluble in NaOH solution (1 in 20).

Uses—*Pharmaceutic aid* (tablet diluent, adsorbent or suspending agent).

Cetyl Alcohol—page 1312.

Cholesterol

Cholest-5-en-3-ol, (3β)-, Cholesterin

Cholest-5-en-3β-ol [57-88-5] $C_{27}H_{46}O$ (386.66).

For the structural formula, see page 389.

A steroid alcohol widely distributed in the animal organism. In addition to cholesterol and its esters, several closely related steroid alcohols occur in the yolk of eggs, the brain, milk, fish oils, wool fat

(10 to 20%), etc. These closely resemble it in properties. One of the methods of commercial production involves extraction of it from the unsaponifiable matter in the spinal cord of cattle, using petroleum benzin. Wool fat also is used as a source.

Description White or faintly yellow, almost odorless pearly leaflets or granules; usually acquires a yellow to pale tan color on prolonged exposure to light or to elevated temperatures; melts 147 to 150°.

Solubility—Insoluble in water; 1 g slowly dissolves in 100 mL of alcohol or about 50 mL of dehydrated alcohol; soluble in acetone, hot alcohol, chloroform, dioxane, ether, ethyl acetate, solvent hexane or vegetable oils.

Uses—To enhance incorporation and emulsification of medicinal products in oils or fats. It is a *pharmaceutical necessity* for *Hydrophilic Petrolatum*, in which it enhances water-absorbing capacity. See Chapter 19.

Dioctyl Sodium Sulfosuccinate (Docusate Sodium)—page 789.

Gelatin

White Gelatin

A product obtained by the partial hydrolysis of collagen derived from the skin, white connective tissues and bones of animals. Gelatin derived from an acid-treated precursor is known as Type A and exhibits an isoelectric point between pH 7 and 9, while gelatin derived from an alkali-treated precursor is known as Type B and exhibits an isoelectric point between pH 4.7 and 5.2.

Gelatin for use in the manufacture of capsules in which to dispense medicines, or for the coating of tablets, may be colored with a certified color, may contain not more than 0.15% of sulfur dioxide, may contain a suitable concentration of sodium lauryl sulfate and suitable antimicrobial agents, and may have any suitable gel strength that is designated by Bloom Gelometer number.

Regarding the special gelatin for use in the preparation of emulsions, see *Emulsions* (page 1534).

Description—Sheets, flakes or shreds, or a coarse to fine powder; faintly yellow or amber in color, the color varying in depth according to the particle size; slight, characteristic bouillon-like odor; stable in air when dry, but is subject to microbial decomposition when moist or in solution.

Solubility—Insoluble in cold water, but swells and softens when immersed in it, gradually absorbing from 5 to 10 times its own weight of water; soluble in hot water, acetic acid or hot mixtures of glycerin and water; insoluble in alcohol, chloroform, ether or fixed and volatile oils.

Uses—In pharmacy, to coat pills and form capsules, and as a vehicle for suppositories. It also is recommended as an emulsifying agent. See under *Emulsions* in Chapters 19 and 83, also *Suppositories* (page 1669), and *Absorbable Gelatin Sponge* (page 816). It also has been used as an adjuvant protein food in malnutrition.

Glyceryl Monostearate—page 1312.

Hydroxyethyl Cellulose

Cellulose, 2-hydroxyethyl ether; Cellosize (*Union Carbide*); Natrosol (*Hercules*)

Cellulose hydroxyethyl ether [9004-62-0].

Preparation—Cellulose is treated with NaOH and then reacted with ethylene oxide.

Description—White, odorless, tasteless, free-flowing powder; softens at about 137°; refractive index (2% solution) about 1.336; pH about 7; solutions are nonionic.

Solubility—Dissolves readily in cold or hot water to give clear, smooth, viscous solutions; partially soluble in acetic acid; insoluble in most organic solvents.

Uses—Resembles carboxymethylcellulose sodium in that it is a cellulose ether, but differs in being nonionic and, hence, its solutions are unaffected by cations. It is used pharmaceutically as a thickener, protective colloid, binder, stabilizer and suspending agent in emulsions, jellies and ointments, lotions, ophthalmic solutions, suppositories and tablets.

Hydroxypropyl Cellulose

Cellulose, 2-hydroxypropyl ether; Klucel (*Hercules*)

Cellulose hydroxypropyl ether [9004-64-2].

Preparation—After treating with NaOH, cellulose is reacted with propylene oxide at elevated temperature and pressure.

Description—Off-white, odorless, tasteless powder; softens at 130°; burns out completely about 475° in N_2 or O_2; refractive index (2% solution) about 1.337; pH (aqueous solution) 6 to 8.5; solutions are nonionic.

Solubility—Soluble in water below 40° (insoluble above 45°); soluble in many polar organic solvents.

Uses—A broad combination of properties useful in a variety of industries. It is used pharmaceutically as a binder, granulation agent and film-coater in the manufacture of tablets; an alcohol-soluble thickener and suspending agent for elixirs and lotions and a stabilizer for emulsions.

Hydroxypropyl Methylcellulose

Cellulose, 2-hydroxypropyl methyl ether

Cellulose hydroxypropyl methyl ether [9004-65-3], available in grades containing 16.5 to 30.0% of methoxy and 4.0 to 32.0% of hydroxypropoxy groups, and thus in viscosity and thermal gelation temperatures of solutions of specified concentration.

Preparation—The appropriate grade of methylcellulose (see below) is treated with NaOH and reacted with propylene oxide at elevated temperature and pressure and for a reaction time sufficient to produce the desired degree of attachment of methyl and hydroxypropyl groups by ether linkages to the anhydroglucose rings of cellulose.

Description—White to slightly off-white, fibrous or granular, free-flowing powder.

Solubility—Swells in water and produces a clear to opalescent, viscous colloidal mixture; undergoes reversible transformation from sol to gel on heating and cooling, respectively. Insoluble in anhydrous alcohol, ether or chloroform.

Uses—A protective colloid that is useful as a dispersing and thickening agent, and in ophthalmic solutions to provide the demulcent action and viscous properties essential for contact-lens use and in "artificial-tear" formulations. See *Hydroxypropyl Methylcellulose Ophthalmic Solution* (page 760).

Lanolin, Anhydrous—page 1311.

Methylcellulose

Cellulose, methyl ether; Methocel (Dow); Cellathyl (*Warner Chilcott*); Hydrolose (*Upjohn*); Syncelose (*Blue Line*)

Cellulose methyl ether [9004-67-5]; a methyl ether of cellulose containing 27.5 to 31.5% of methoxy groups.

Preparation—By the reaction of methyl chloride or of dimethyl sulfate on cellulose dissolved in sodium hydroxide. The cellulose methyl ether so formed is coagulated by adding methanol or other suitable agent and centrifuged. Since cellulose has 3 hydroxyl groups/glucose residue, several methylcelluloses can be made varying, among other properties, in solubility and viscosity. Types useful for pharmaceutical application contain from 1 to 2 methoxy radicals/glucose residue.

Description—White, fibrous powder or granules; aqueous suspensions neutral to litmus; stable to alkalies and dilute acids.

Solubility—Insoluble in ether, alcohol or chloroform; soluble in glacial acetic acid and in a mixture of equal parts of alcohol and chloroform; swells in water, producing a clear to opalescent, viscous colloidal solution; insoluble in hot water and saturated salt solutions; salts of minerals acids and particularly of polybasic acids, phenols and tannins coagulate its solutions, but this can be prevented by the addition of alcohol or of glycol diacetate.

Uses—A synthetic substitute for natural gums that has both pharmaceutic and therapeutic applications. Pharmaceutically, it is used as a *dispersing, thickening, emulsifying, sizing* and *coating agent*. It is an ingredient of many nose drops, eye preparations, burn medications, cosmetics, tooth pastes, liquid dentifrices, hair fixatives, creams and lotions. It functions as a protective colloid for

many types of dispersed substances and is an effective stabilizer for oil-in-water emulsions.

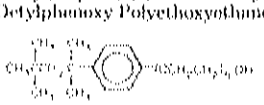
Therapeutically, it is used as a *bulk laxative* in the treatment of *chronic constipation*. Taken with 1 or 2 glasseful of water, it forms a colloidal solution in the upper alimentary tract; this solution loses water in the colon, forming a gel that increases the bulk and softness of the stool. The gel is bland, demulcent and nonirritating to the gastrointestinal tract. Once a normal stool develops, the dose should be reduced to a level adequate for maintenance of good function. Although it takes up water from the gastrointestinal tract quite readily, methylcellulose tablets have caused fecal impaction and intestinal obstruction when taken with a limited amount of water. It also is used as a topical ophthalmic protectant, in the form of 0.5 to 1% solution serving as artificial tears or a contact-lens solution applied to the conjunctiva, 0.05 to 0.1 mL at a time, 3 or 4 times a day as needed.

Dose—*Usual*, as laxative, 1 to 1.5 g, with water, 2 to 4 times a day.

Dosage Forms Tablets: 500 mg; Ophthalmic Solution: 0.5 and 1%; Syrup: 5.91 g/30 mL.

Octoxynol 9

Polyoxy-1,2-ethanediyl, α-[4-(1,1,3,3-tetramethylbutyl)phenyl]-ω-hydroxy-, Octylphenoxy Polyethoxyethanol NF XII



Polyethylene glycol mono[*p*-(1,1,3,3-tetramethylbutyl)phenyl] ether [9002-93-1]; an anhydrous liquid mixture of mono-*p*-(1,1,3,3-tetramethylbutyl)phenyl ethers of polyethylene glycols in which *n* varies from 5 to 15, and which has an average molecular weight of 647, corresponding to the formula $C_{34}H_{62}O_{11}$.

Preparation—By reacting *p*-(1,1,3,3-tetramethylbutyl)phenol with ethylene oxide at elevated temperature under pressure in the presence of NaOH.

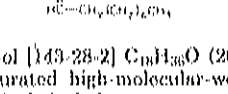
Description—Clear, pale yellow, viscous liquid, having a faint odor and a bitter taste; specific gravity about 1.064; pH (1 in 100 aqueous solution) about 7.

Solubility—Miscible with water, alcohol or acetone; soluble in benzene or toluene; insoluble in solvent hexane.

Uses—A nonionic detergent, emulsifier and dispersing agent. It is an ingredient in *Nitrofurazone Solution*. See *Polyethylene Glycol 400* (page 1313).

Oleyl Alcohol

9-Octadecen-1-ol, (*Z*)-, Adol 85 (*Sherex*)



(*Z*)-9-Octadecen-1-ol [143-28-2] $C_{18}H_{36}O$ (268.48); a mixture of unsaturated and saturated high-molecular-weight fatty alcohols consisting chiefly of oleyl alcohol.

Preparation—One method reacts ethyl oleate with absolute ethanol and metallic sodium (*Org Syn Coll III*: 673, 1955).

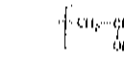
Description—Clear, colorless to light yellow, oily liquid; faint characteristic odor and bland taste; iodine value between 85 and 90; hydroxyl value between 205 and 215.

Solubility—Soluble in alcohol, ether, isopropyl alcohol or light mineral oil; insoluble in water.

Uses—*A pharmaceutic aid* (emulsifying agent or emollient).

Polyvinyl Alcohol

Ethenol, homopolymer



Vinyl alcohol polymer [9002-89-5] $(C_2H_4O)_n$.

Preparation—Polyvinyl acetate is approximately 88% hydrolyzed in a methanol-methyl acetate solution using either mineral acid or alkali as a catalyst.

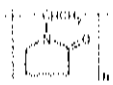
Description—White to cream-colored powder or granules; odorless.

Solubility—Freely soluble in water; solution effected more rapidly at somewhat elevated temperatures.

Uses—A *suspending agent* and *emulsifier*, either with or without the aid of a surfactant. It commonly is employed as a lubricant and protectant in various ophthalmic preparations, such as decongestants, artificial tears and contact-lens products (see page 1593).

Povidone

2-Pyrrolidinone, 1-ethenyl-, homopolymer; Polyvinylpyrrolidone; PVP



1-Vinyl-2-pyrrolidinone polymer [9003-39-8] $(C_6H_9NO)_n$; a synthetic polymer consisting of linear 1-vinyl-2-pyrrolidinone groups, the degree of polymerization of which results in polymers of various molecular weights. It is produced commercially as a series of products having mean molecular weights ranging from about 10,000 to about 700,000. The viscosity of solutions containing 10% or less is essentially the same as that of water; solutions more concentrated than 10% become more viscous, depending upon the concentration and the molecular weight of the polymer used. It contains 12 to 13% of nitrogen.

Preparation—1,4-Butanediol is dehydrogenated thermally with the aid of copper to γ-butyrolactone, which is then reacted with ammonia to form 2-pyrrolidinone. Addition of the latter to acetylene yields vinylpyrrolidinone (monomer) which is polymerized thermally in the presence of hydrogen peroxide and ammonia.

Description—White to creamy white, odorless powder, hygroscopic; pH (1 in 20 solution) 3 to 7.

Solubility—Soluble in water, alcohol or chloroform; insoluble in ether.

Uses—A *dispersing* and *suspending* agent in pharmaceutical preparations.

Propylene Glycol Monostearate

Octadecanoic acid, monoester with 1,2-propanediol

1,2-Propanediol monostearate [1323-39-3]; a mixture of the propylene glycol mono- and diesters of stearic and palmitic acids. It contains not less than 90% of monoesters of saturated fatty acids, chiefly propylene glycol monostearate ($C_{21}H_{42}O_3$) and propylene glycol monopalmitate ($C_{19}H_{38}O_3$).

Preparation—By reacting propylene glycol with stearoyl chloride in a suitable dehydrochlorinating environment.

Description—White, wax-like solid or white, wax-like beads or flakes; slight, agreeable, fatty odor and taste; congeals not lower than 45°; acid value not more than 2; saponification value 155 to 165; hydroxyl value 150 to 170; iodine value not more than 3.

Solubility—Dissolves in organic solvents such as alcohol, mineral or fixed oils, benzene, ether or acetone; insoluble in water but may be dispersed in hot water with the aid of a small amount of soap or other suitable surface-active agent.

Uses—A *surfactant*. It is particularly useful as a dispersing agent for perfume oils or oil-soluble vitamins in water, and in cosmetic preparations.

Silicon Dioxide, Colloidal—page 1325.

Sodium Lauryl Sulfate

Sulfuric acid monododecyl ester sodium salt; Irium; Duponol C (*Dupont*); Gardinol WA (*Procter & Gamble*)

Sodium monododecyl sulfate [151-21-3]; a mixture of sodium alkyl sulfates consisting chiefly of sodium lauryl sulfate. The combined content of sodium chloride and sodium sulfate is not more than 8%.

Preparation—The fatty acids of coconut oil, consisting chiefly of lauric acid, are catalytically hydrogenated to form the corresponding alcohols. The latter are then esterified with sulfuric acid (sulfated) and the resulting mixture of alkyl bisulfates (alkylsulfuric acids) is converted into a mixture of sodium salts by reacting with alkali under controlled conditions of pH.

Description—Small, white or light yellow crystals having a slight, characteristic odor.

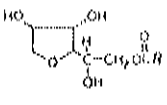
Solubility—1 g in 10 mL water, forming an opalescent solution.

Incompatibilities—Reacts with *cationic surface-active agents* with loss of activity, even in concentrations too low to cause precipitation. Unlike soaps, it is compatible with dilute acids, and calcium and magnesium ions.

Uses—An emulsifying, detergent and wetting agent in ointments, tooth powders and other pharmaceutical preparations, and in the metal, paper and pigment industries. See Chapters 19 and 87.

Sorbitan Esters

Spans (*Atlas*)



Sorbitan esters (*monolaurate* [1338-39-2]; *monooleate* [1338-43-8]; *monopalmitate* [26266-57-9]; *monostearate* [1338-41-6]; *trioleate* [26266-58-0]; *tristearate* [26658-19-5]).

Preparation—Sorbitol is dehydrated to form a *hexitan* which is then esterified with the desired fatty acid. See *Polysorbates*, page 1314, which are polyethylene glycol ethers of sorbitan fatty acid esters.

Description—*Monolaurate:* Amber, oily liquid; may become hazy or form a precipitate; viscosity about 4250 cps; HLB no 8.6; acid no 7.0 max; saponification no 158 to 170; hydroxyl no 330 to 358. *Monooleate:* Amber liquid; viscosity about 1000 cps; HLB no 4.3; acid no 8.0 max; saponification no 145 to 160; hydroxyl no 193 to 210. *Monopalmitate:* Tan, granular waxy solid; HLB no 6.7; acid no 4 to 7.5; saponification no 140 to 150; hydroxyl no 275 to 305. *Monostearate:* Cream to tan beads; HLB no 4.7; acid no 5 to 10; saponification no 147 to 157; hydroxyl no 235 to 260. *Trioleate:* Amber, oily liquid; viscosity about 200 cps; HLB no 1.8; acid no 15 max; saponification no 170 to 190; hydroxyl no 65 to 70. *Tristearate:* Tan, waxy beads; HLB no 2.1; acid no 12 to 15; saponification no 176 to 188; hydroxyl no 66 to 80.

Solubility—*Monolaurate:* Soluble in methanol or alcohol; dispersible in distilled water and hard water (200 ppm); insoluble in hard water (20,000 ppm). *Monooleate:* Soluble in most mineral or vegetable oils; slightly soluble in ether; dispersible in water; insoluble in acetone. *Monopalmitate:* Dispersible (50°) in distilled water or hard water (200 ppm); soluble in ethyl acetate; insoluble in cold distilled water or hard water (20,000 ppm). *Monostearate:* Soluble (above melting point) in vegetable oils or mineral oil; insoluble in water, alcohol and propylene glycol. *Trioleate:* Soluble in mineral oil, vegetable oils, alcohol or methanol; insoluble in water. *Tristearate:* Soluble in isopropyl alcohol; insoluble in water.

Uses—Nonionic *surfactants* used as *emulsifying agents* in the preparation of water-in-oil emulsions.

Stearic Acid—page 1312.

Stearyl Alcohol

1-Octadecanol [112-92-5] $C_{18}H_{38}O$ (270.50); contains not less than 90% of stearyl alcohol, the remainder consisting chiefly of cetyl alcohol [$C_{16}H_{34}O$ = 242.44].

Preparation—Through the reducing action of lithium aluminum hydride on ethyl stearate.

Description—White, unctuous flakes or granules having a faint, characteristic odor and a bland taste; melts 55 to 60°.

Solubility—Insoluble in water; soluble in alcohol, chloroform, ether or vegetable oils.

Uses—A surface-active agent used to *stabilize emulsions* and increase their ability to retain large quantities of water. See *Hydrophilic Ointment* (page 1312), *Hydrophilic Petrolatum* (page 1311), and Chapters 19 and 87.

Sterculia Gum—page 780.

Tragacanth

Gum Tragacanth; Hog Gum; Goat's Thorn

The dried gummy exudation from *Astragalus gummifer* Labillardière, or other Asiatic species of *Astragalus* (Fam. *Leguminosae*).

Constituents—60 to 70% bassorin and 30 to 40% soluble gum (*tragacanthin*). The bassorin swells in the presence of water to form a gel and tragacanthin forms a colloidal solution. Bassorin, consisting of complex methoxylated acids, resembles pectin. Tragacanthin yields glucuronic acid and arabinose when hydrolyzed.

Description—Flattened, lamellated, frequently curved fragments or straight or spirally twisted linear pieces 0.5 to 2.5 mm in thickness; white to weak-yellow in color; translucent; horny in texture; odorless; insipid, mucilaginous taste. When powdered, it is white to yellowish white.

Introduced into water, tragacanth absorbs a certain proportion of that liquid, swells very much, and forms a soft adhesive paste, but does not dissolve. If agitated with an excess of water, this paste forms a uniform mixture; but in the course of 1 or 2 days the greater part separates, and is deposited, leaving a portion dissolved in the supernatant fluid. The finest mucilage is obtained from the whole gum or *flake* form. Several days should be allowed for obtaining a uniform mucilage of the maximum gel strength. A common adulterant is *Karaya Gum*, and the USP/NF has introduced tests to detect its presence.

Solubility—Insoluble in alcohol.

Uses—A *suspending agent* in lotions, mixtures and extemporaneous preparations and prescriptions. It is used with emulsifying agents largely to increase consistency and retard creaming. It is sometimes used as a *demulcent* in sore throat, and the jelly-like product formed when the gum is allowed to swell in water serves as a basis for pharmaceutical jellies, eg, *Ephedrine Sulfate Jelly*. It also is used in various confectionery products. In the form of a glycerite, it has been used as a pill excipient.

Tragacanth Mucilage—*Preparation:* Mix glycerin (18 g) with purified water (75 mL) in a tared vessel, heat the mixture to boiling, discontinue the application of heat, add tragacanth (6 g) and benzoic acid (0.2 g) and macerate the mixture during 24 hr, stirring occasionally. Then add enough purified water to make the mixture weigh 100 g, stir actively until of uniform consistency, and strain forcibly through muslin. *Uses:* A suspending agent for insoluble substances in internal mixtures. It is also a *protective* agent.

Xanthan Gum

Keltrol (*Kelco*)

A high-molecular-weight polysaccharide gum produced by a pure-culture fermentation of a carbohydrate with *Xanthomonas campestris*, then purified by recovery with isopropyl alcohol, dried and milled; contains D-glucose and D-mannose as the dominant hexose units, along with D-glucuronic acid, and is prepared as sodium, potassium or calcium salt; yields 4.2 to 5% of carbon dioxide.

Preparation—See above and US Pats 3,433,708 and 3,557,016.

Description—White or cream-colored, tasteless powder with a slight organic odor; powder and solutions stable at 25° or less; does not exhibit polymorphism; aqueous solutions are neutral to litmus.

Solubility—1 g in about 3 mL of alcohol; soluble in hot or cold water.

Uses—A hydrophilic colloid to thicken, suspend, emulsify and stabilize water-based systems.

Other Emulsifying and Suspending Agents

Chondrus [Irish Moss; Carrageenan]—The dried sun-bleached plant of *Chondrus crispus* (Linné) Stackhouse (Fam *Gigartinaceae*). *Uses:* Principally, as an emulsifying agent for liquid petrolatum and for cod liver oil. It is also a protective.

Malt—The partially germinated grain of one or more varieties of *Hordeum vulgare* Linné (Fam *Gramineae*) and contains amylolytic enzymes. Yellowish or amber-colored grains, having a characteristic odor and a sweet taste. The evaporated aqueous extract constitutes malt extract.

Malt Extract—The product obtained by extracting malt, the partially and artificially germinated grain of one or more varieties of *Hordeum vulgare* Linné (Fam *Gramineae*). *Uses:* An infrequently used emulsifying agent.

Ointment Bases

Ointments are semisolid preparations for external application to the body. They should be of such composition that they soften, but not necessarily melt, when applied to the skin. Therapeutically, ointments function as protectives and emollients for the skin, but are used primarily as vehicles or bases for the topical application of medicinal substances. Ointments also may be applied to the eye or eyelids.

Ideally, an ointment base should be compatible with the skin, stable, permanent, smooth and pliable, nonirritating, nonsensitizing, inert and readily able to release its incorporated medication. Since there is no single ointment base which possesses all these characteristics, continued research in this field has resulted in the development of numerous new bases. Indeed, ointment bases have become so numerous as to require classification. Although ointment bases may be grouped in several ways, it is generally agreed that they can be classified best according to composition. Hence, the following four classes are recognized herein: oleaginous, emulsifiable, emulsion bases and water-soluble.

For completeness, substances are included that, although not used alone as ointment bases, contribute some pharmaceutical property to one or more of the various bases.

Oleaginous Ointment Bases and Components

The oleaginous ointment bases include fixed oils of vegetable origin, fats obtained from animals and semisolid hydrocarbons obtained from petroleum. The vegetable oils are used chiefly in ointments to lower the melting point or to soften bases. These oils can be used as a base in themselves when a high percentage of powder is incorporated.

The vegetable oils and the animal fats have two marked disadvantages as ointment bases: their water-absorbing capacity is low and they have a tendency to become rancid. Insofar as vegetable oils are concerned, the second disadvantage can be overcome by hydrogenation, a process which converts many fixed oils into white, semisolid fats or into hard, almost brittle, waxes.

The hydrocarbon bases comprise a group of substances with a wide range of melting points so that any desired consistency and melting point may be prepared with representatives of this group. They are stable, bland, chemically inert and will mix with virtually any chemical substance. Oleaginous bases are excellent emollients.

White Ointment

Ointment USP XI; Simple Ointment

White Wax	50 g
White Petrolatum	950 g
To make	1000 g

Melt the white wax in a suitable dish on a water bath, add the white petrolatum, warm until liquefied, then discontinue the heating, and stir the mixture until it begins to congeal. It is permissible to vary the proportion of wax to obtain a suitable consistency of the ointment under different climatic conditions.

Uses—An emollient and vehicle for other ointments.

Yellow Ointment

Yellow Wax	50 g
Petrolatum	950 g
To make	1000 g

Melt the yellow wax in a suitable dish on a steam bath, add the petrolatum, warm until liquefied, then discontinue the heating, and stir the mixture until it begins to congeal. It is permissible to vary the proportion of wax to obtain a suitable consistency of the ointment under different climatic conditions.

Uses—An emollient and vehicle for other ointments. Both white and yellow ointment are known as "simple ointment." White ointment should be used to prepare white ointments and yellow ointments should be used to prepare colored ointments when simple ointment is prescribed.

Cetyl Esters Wax

"Synthetic Spermaceti"

A mixture consisting primarily of esters of saturated fatty alcohols (C_{14} to C_{18}) and saturated fatty acids (C_{14} to C_{18}). It has a saponification value of 109 to 120 and an acid value of not more than 5.

Description—White to off-white, somewhat translucent flakes; crystalline structure and pearly luster when caked; faint odor and a bland, mild taste; free from rancidity; specific gravity 0.820 to 0.840 at 50°; iodine value not more than 1; melts 43 to 47°.

Solubility—Insoluble in water; practically insoluble in cold alcohol; soluble in boiling alcohol, ether, chloroform or fixed and volatile oils; slightly soluble in cold solvent hexane.

Uses—A replacement for spermaceti used to give consistency and texture to ointments, eg, *Cold Cream* and *Rose Water Ointment*.

Oleic Acid

(*Z*)-9-Octadecenoic acid; Oleinic Acid; Elaic Acid

$$\begin{matrix} HC{-}CH_2(CH_2)_6COOH \\ \| \\ HC{-}CH_2(CH_2)_6CH_3 \end{matrix}$$

Oleic acid [112-80-1] obtained from tallow and other fats, and consists chiefly of (*Z*)-9-octadecenoic acid (282.47). Oleic acid used in preparations for internal administration is derived from edible sources.

It usually contains variable amounts of the other fatty acids present in tallow such as linolenic and stearic acids.

Preparation—Obtained as a by-product in the manufacture of the solid stearic and palmitic acids used in the manufacture of candles, stearates and other products. The crude oleic acid is known as "red oil," the stearic and palmitic acids being separated by cooling.

Description—Colorless to pale yellow, oily liquid; lard-like odor and taste; specific gravity 0.889 to 0.896; congeals at a temperature not above 10°; pure acid solidifies at 4°; at atmospheric pressure it decomposes when heated at 80 to 100°; on exposure to air it gradually absorbs oxygen, darkens and develops a rancid odor.

Solubility—Practically insoluble in water; miscible with alcohol, chloroform, ether, benzene or fixed and volatile oils.

Incompatibilities—Reacts with *alkalies* to form soaps. *Heavy metals* and *calcium salts* form insoluble oleates. *Iodine solutions* are decolorized by formation of the iodine addition compound of the acid. It is oxidized to various derivatives by *nitric acid*, *potassium permanganate* and other agents.

Uses—Classified as an emulsion adjunct, which reacts with alkalis to form soaps that function as emulsifying agents; it is used for this purpose in such preparations as *Benzyl Benzoate Lotion* and *Green Soap*. It also is used to prepare oleate salts of bases.

Olive Oil

Sweet Oil

The fixed oil obtained from the ripe fruit of *Olea europaea* Linné (Fam *Oleaceae*).

Preparation—By crushing recently collected ripe olives in a mill without breaking the putamen, then moderately pressing the pulpy mass. This produces the highest grade oil, known as *virgin oil*, "sublime oil" or "first expressed oil." The mass in the press then is mixed with water and again expressed with greater pressure, an oil of second quality resulting. Any oil remaining in the press cake is finally extracted with carbon disulfide, or the mass is thrown into large cisterns, mixed with water and the oil allowed to separate. This is sometimes called "Pyrene oil," "bagasse oil" or "huile d'enfer." When bought in bulk or from unlabeled containers, cottonseed oil, colza oil, grapeseed oil, sesame oil or other bland oils are not uncommonly found as adulterants. Large quantities of this oil are imported from Italy and other countries bordering the Mediterranean, and it is produced to a limited extent in the Southern US, chiefly in California.

Description—Pale yellow or light greenish yellow, oily liquid; slight, characteristic odor and taste, with a faintly acrid aftertaste; specific gravity 0.910 to 0.915.
Solubility—Slightly soluble in alcohol; miscible with carbon disulfide, chloroform or ether.

Uses—In making cerates, ointments, liniments, and plasters. It is a bland oil, well-suited for *emollient* purposes and for food. It also is used as an emollient laxative; sufficient must be given so that enough escapes digestion to soften the stool.
Dose—The usual dose is **30 mL.**

Paraffin

Paraffin Wax, Hard Paraffin

A purified mixture of solid hydrocarbons obtained from petroleum.

Description—Colorless or white, more or less translucent mass with a crystalline structure; slightly greasy to the touch; odorless and tasteless; congeals 47 to 65°.
Solubility—Freely soluble in chloroform, ether, volatile oils or most warm fixed oils; slightly soluble in dehydrated alcohol; insoluble in water or alcohol.

Uses—Mainly, to increase the consistency of some ointments.

Petrolatum

Yellow Soft Paraffin; Amber Petrolatum; Yellow Petrolatum; Petroleum Jelly; Paraffin Jelly

A purified mixture of semisolid hydrocarbons obtained from petroleum. It may contain a suitable stabilizer.
Preparation—The "residuums," as they are termed technically, which are obtained by the distillation of petroleum, are purified by melting, usually treating with sulfuric acid and then percolating through recently burned bone black or adsorptive clays; this removes the odor and modifies the color. Selective solvents are also sometimes employed to extract impurities.
It has been found that the extent of purification required to produce *Petrolatum* and *Light Mineral Oil* of official quality removes antioxidants that are naturally present, and the purified product subsequently has a tendency to oxidize and develop an offensive odor. This is prevented by the addition of a minute quantity of α-tocopherol, or other suitable antioxidant, as is now permissible.

Description—Unctuous mass of yellowish to light amber color; not more than a slight fluorescence after being melted; transparent in thin layers; free or nearly free from odor and taste; specific gravity 0.815 to 0.880 at 60°; melts between 38 and 60°.
Solubility—Insoluble in water; almost insoluble in cold or hot alcohol or in cold dehydrated alcohol; freely soluble in benzene, carbon disulfide, chloroform or turpentine oil; soluble in ether, solvent hexane or in most fixed and volatile oils, the degree of solubility in these solvents varying with the composition of the petrolatum.

Uses—A base for ointments. It is highly occlusive and therefore a good emollient but it may not release some drugs readily.

White Petrolatum

White Petroleum Jelly; White Soft Paraffin

A purified mixture of semisolid hydrocarbons obtained from petroleum, and wholly or nearly decolorized. It may contain a suitable stabilizer.
Preparation—In the same manner as petrolatum, the purification treatment being continued until the product is practically free from yellow color.

Description—White or faintly yellowish, unctuous mass; transparent in thin layers, even after cooling to 0°; specific gravity 0.815 to 0.880 at 60°; melts 38 to 60°.
Solubility—Similar to that described under *Petrolatum*.

Uses—Similar to yellow petrolatum but often is preferred because of its freedom from color. It is employed as a protective, a base for ointments and cerates and to form the basis for burn dressings. See *Petrolatum Gauze* (page 758).

Spermaceti

A waxy substance obtained from the head of the sperm whale, *Physeter macrocephalus* Linné (Fam *Physeteridae*).
Constituents—A mixture of several constituents of which cetin, or cetyl palmitate [$C_{15}H_{31}COOC_{16}H_{33}$], predominates. When recrystallized from alcohol, *cetin* is obtained, while the mother liquor on evaporation deposits an oil, *cetin elain*, which when saponified yields *cetin elaic acid*, an acid resembling, but distinct from, oleic acid.
Preparation—By pumping the oleaginous material from the head of the sperm whale, separating the liquid portion known as sperm oil and purifying the remaining crude solid, which is this substance.

Description—White, somewhat translucent, slightly unctuous masses; crystalline fracture and pearly luster; faint odor and a bland, mild taste; free from rancidity; specific gravity about 0.94; melts 44 to 52°.
Solubility—Insoluble in water; practically insoluble in cold alcohol; slightly soluble in cold solvent hexane; soluble in boiling alcohol, ether, chloroform or fixed and volatile oils.

Uses—One of the solid fatty substances formerly employed to give consistency and texture to cerates and ointments, as in *Cold Cream* and *Rose Water Ointment*. In the interest of whale conservation, this has been replaced by *cetyl esters wax* (also known as *synthetic spermaceti*).
Dose—*For external use, topically*, as required.

Starch Glycerite

Starch Glycerin

Starch	100 g
Benzoic Acid	2 g
Purified Water	200 mL
Glycerin	700 mL
To make about	1000 g

Rub the starch and the benzoic acid with the purified water in a porcelain dish until a smooth mixture is produced, then add the glycerin, and mix well. Heat the mixture on a sand bath to a temperature between 140 and 144°, with constant but gentle stirring until a translucent, jelly-like mass results, and then strain through muslin.
It should be freshly prepared.

Uses—Although not an oleaginous base, this *emollient* preparation is sometimes used as a substitute for a fatty ointment. It also has been used as a *pill excipient*.
Dose—*For external use, topically*, as required.

White Wax

Bleached Beeswax; White Beeswax; Bleached Wax

The product of bleaching and purifying yellow wax that is obtained from the honeycomb of the bee [*Apis mellifera* Linné (Fam *Apidae*)].
Preparation—The color of yellow wax is discharged by exposing it with an extended surface to the combined influence of air, light and moisture. In one process a stream of melted wax is directed on a revolving cylinder kept constantly wet, upon which it congeals in thin layers which are spread on linen cloths stretched on frames and

exposed to the air and light, care being taken to wet and occasionally turn them. In a few days they are partially bleached; but to remove the color completely it is necessary to repeat the whole process one or more times. When sufficiently bleached, it is melted and cast into small circular cakes.

Description—Yellowish white, nearly tasteless, somewhat translucent solid; faint, characteristic odor; free from rancidity; melts 62 to 65°; specific gravity about 0.95.

Solubility—Insoluble in water; sparingly soluble in cold alcohol; boiling alcohol dissolves the cerotic acid and a portion of the myricin, which are constituents; completely soluble in chloroform, ether or fixed and volatile oils; partly soluble in cold benzene or cold carbon disulfide; completely soluble in these liquids at about 30°.

Uses—A stiffening agent in many preparations such as cerates, pastes and ointments.

Yellow Wax

Beeswax; Yellow Beeswax

The purified wax from the honeycomb of the bee, *Apis mellifera* Linné (Fam *Apidae*).

Constituents—A mixture of three substances: (1) *myricin*, insoluble in boiling alcohol and consisting chiefly of *myricyl palmitate* [$C_{30}H_{61}(C_{16}H_{31}O_2)$] and *myricyl alcohol* [$C_{30}H_{61}OH$]; (2) *cerin* or *cerotic acid* [$C_{26}H_{52}O_2$], formerly called *cerin* when obtained only in an impure state, which is dissolved by boiling alcohol, but crystallizes out on cooling and (3) *cerolein*, which remains dissolved in the cold alcoholic liquid. The latter is probably a mixture of fatty acids.

Preparation—It is a natural secretion of bees. It is obtained on the large scale by first abstracting the honey from the combs by shaving off the ends of the cells, draining and then placing them in centrifuges. The honey is rapidly whirled out, water is added and the wax is cleaned thoroughly and quickly; it then is melted and strained and run into molds to cool and harden.

Description—Yellow to grayish brown solid; agreeable, honeylike odor; faint, characteristic taste; when cold it is somewhat brittle and when broken it presents a dull, granular, noncrystalline fracture; becomes pliable from the heat of the hand; specific gravity about 0.95; melts between 62 and 65°.

Solubility—Insoluble in water; sparingly soluble in cold alcohol; completely soluble in chloroform, ether or fixed and volatile oils; partly soluble in cold benzene or carbon disulfide; completely soluble in these liquids at about 30°.

Uses—A stiffening agent in many pharmaceutical preparations and ingredient of many polishes.

Absorbent Ointment Bases

The term absorbent is used here to denote the water-absorbing or emulsifying properties of these bases and not to describe their action on the skin. These bases, sometimes called *emulsifiable ointment bases*, are generally anhydrous substances which have the property of absorbing (emulsifying) considerable quantities of water and still retaining their ointment-like consistency. Preparations of this type do not contain water as a component of their basic formula, but if water is incorporated, when and as desired, a W/O emulsion results. The following official products fall into this category.

Anhydrous Lanolin

Wool Fat USP XVI; Refined Wool Fat

Lanolin that contains not more than 0.25% of water.

Constituents—Contains the sterols *cholesterol* [$C_{27}H_{45}OH$] and *oxycholesterol*, as well as triterpene and aliphatic alcohols. About 7% of the alcohols are found in the free state, the remainder occurring as esters of the following fatty acids: *carnaubic*, *cerotic*, *lanoceric*, *lanopalmitic*, *myristic* and *palmitic*. Some of these are found free. The emulsifying and emollient actions of lanolin are due to the alcohols that are found in the unsaponifiable fraction when lanolin is treated with alkali. Constituting approximately one-half of this fraction and known as *lanolin alcohols*, the latter is comprised of *cholesterol* (30%), *lanosterol* (25%), *cholestanol* (*dihydrocholesterol*) (3%), *agnosterol* (2%) and various other alcohols (40%).

Preparation—By purifying the fatty matter (*suint*) obtained from the wool of the sheep. This natural wool fat contains about 30% of free fatty acids and fatty acid esters of *cholesterol* and other higher alcohols. The cholesterol compounds are the important constituents and, to secure these in a purified form, many processes have been devised. In one of these the crude wool fat is treated with weak alkali, the saponified fats and emulsions centrifuged to secure the aqueous soap solution, from which, on standing, a layer of partially purified wool fat separates. This product is further purified by treating it with calcium chloride and then dehydrated by fusion with unslaked lime. It is finally extracted with acetone and the solvent subsequently separated by distillation. This differs from lanolin in that the former contains practically no water.

Description—Yellow, tenacious, unctuous mass; slight, characteristic odor; melts between 36 and 42°.

Solubility—Insoluble in water, but mixes without separation with about twice its weight of water; sparingly soluble in cold alcohol; more soluble in hot alcohol; freely soluble in ether or chloroform.

Uses—An ingredient of ointments, especially when an aqueous liquid is to be incorporated. It gives a distinctive quality to the ointment, increasing absorption of active ingredients and maintaining a uniform consistency for the ointment under most climatic conditions. However, it has been omitted from many ointments on the recommendation of dermatologists who have found that many patients are allergic to this animal wax.

Hydrophilic Petrolatum

Cholesterol	30 g
Stearyl Alcohol	30 g
White Wax	80 g
White Petrolatum	860 g
To make	1000 g

Melt the stearyl alcohol, white wax, and white petrolatum together on a steam bath, then add the cholesterol, and stir until it completely dissolves. Remove from the bath, and stir until the mixture congeals.

Uses—A *protective* and *water-absorbable ointment base*. It will absorb a large amount of water from aqueous solutions of medicating substances, forming a W/O type of emulsion. See *Ointments* (page 1602).

Other Absorption Ointment Bases

Hydroxystearin Sulfate [Sulfated Hydrogenated Castor Oil; SHCO] A substance prepared by sulfating hydrogenated castor oil. Pale, yellow-brown, unctuous semisolid mass; faint odor containing about 9% organically bound SO_3. Dispersible in water and glycerin; miscible with propylene glycol, petrolatums or fixed oils. *Uses:* A surface-active agent used in preparing hydrophilic ointment bases and other emulsions.

Emulsion Ointment Bases and Components

Emulsion ointment bases are actually semisolid emulsions. These preparations can be divided into two groups on the basis of emulsion type: emulsion ointment base water-in-oil (W/O) type and emulsion ointment base oil-in-water

(O/W) type. Bases of both types will permit the incorporation of some additional amounts of water without reducing the consistency of the base below that of a soft cream. However, only O/W emulsion ointment bases can be removed readily from the skin and clothing with water. W/O emulsions are better emollients and protectants than are O/W emulsions. W/O emulsions can be diluted with oils.

Cetyl Alcohol

Cetostearyl Alcohol; "Palmityl" Alcohol; Aldol 62 (*Sherex*)

$CH_3(CH_2)_{14}CH_2OH$

1-Hexadecanol [124-29-8] $C_{16}H_{34}O$ (242.44); a mixture of not less than 90% of cetyl alcohol, the remainder chiefly stearyl alcohol.

Preparation—By catalytic hydrogenation of palmitic acid, or saponification of spermaceti, which contains cetyl palmitate.

Description—Unctuous, white flakes, granules, cubes or castings; faint characteristic odor and a bland, mild taste; melts 45 to 50°; not less than 90% distils between 316 and 336°.

Solubility—Insoluble in water; soluble in alcohol, chloroform, ether or vegetable oils.

Uses—Similar to *Stearyl Alcohol* (page 1308). It also imparts a smooth texture to the skin and is used widely in cosmetic creams and lotions.

Cold Cream

Petrolatum Rose Water Ointment USP XVI

Cetyl Esters Wax	125 g
White Wax	120 g
Mineral Oil	560 g
Sodium Borate	5 g
Purified Water	190 mL
To make about	1000 g

Reduce the cetyl esters wax and the white wax to small pieces, melt them on a steam bath with the mineral oil and continue heating until the temperature of the mixture reaches 70°. Dissolve the sodium borate in the purified water, warmed to 70° and gradually add the warm solution to the melted mixture, stirring rapidly and continuously until it has congealed.

If the ointment has been chilled, warm it slightly before attempting to incorporate other ingredients (see USP for allowable variations).

Uses—Useful as an emollient, cleansing cream and ointment base. It resembles *Rose Water Ointment*, differing only in that mineral oil is used in place of almond oil and omitting the fragrance. This change produces an ointment base which is not subject to rancidity like one containing a vegetable oil. This is a W/O emulsion.

Glyceryl Monostearate

Octadecanoic acid, monoester with 1,2,3-propanetriol

Monostearin [31566-31-1]; a mixture chiefly of variable proportions of glyceryl monostearate [$C_3H_5(OH)_2C_{18}H_{35}O_2$ = 358.56] and glyceryl monopalmitate [$C_3H_5(OH)_2C_{16}H_{31}O_2$ = 330.51].

Preparation—Among other ways, by reacting glycerin with commercial stearoyl chloride.

Description—White, wax-like solid or occurs in the form of white, wax-like beads, or flakes; slight, agreeable, fatty odor and taste; does not melt below 55°; affected by light.

Solubility—Insoluble in water, but may be dispersed in hot water with the aid of a small amount of soap or other suitable surface-active agent; dissolves in hot organic solvents such as alcohol, mineral or fixed oils, benzene, ether or acetone.

Uses—A thickening and emulsifying agent for ointments. See *Ointments* (page 1602).

Hydrophilic Ointment

Methylparaben	0.25 g
Propylparaben	0.15 g
Sodium Lauryl Sulfate	10 g
Propylene Glycol	120 g
Stearyl Alcohol	250 g
White Petrolatum	250 g
Purified Water	370 g
To make about	1000 g

Melt the stearyl alcohol and the white petrolatum on a steam bath, and warm to about 75°. Add the other ingredients, previously dissolved in the water and warmed to 75°, and stir the mixture until it congeals.

Uses—A *water-removable ointment base* for the so-called "washable" ointments. This is an O/W emulsion.

Lanolin

Hydrous Wool Fat

The purified, fat-like substance from the wool of sheep, *Ovis aries* Linné (Fam *Bovidae*); contains 25 to 30% water.

Description—Yellowish white, ointment-like mass, having a slight, characteristic odor; when heated on a steam bath it separates into an upper oily and a lower water layer; when the water is evaporated a residue of *Lanolin* remains which is transparent when melted.

Solubility—Insoluble in water; soluble in chloroform or ether with separation of its water of hydration.

Uses—Largely as a vehicle for ointments, for which it is admirably adapted, on account of its compatibility with skin lipids. It emulsifies aqueous liquids. Lanolin is a W/O emulsion.

Stearic Acid

Octadecanoic acid; Cetylacetic Acid; Stearophanic Acid

Stearic acid [57-11-4]; a mixture of stearic acid [$C_{18}H_{36}O_2$ = 284.48] and palmitic acid [$C_{16}H_{32}O_2$ = 256.43], which together constitute not less than 90.0% of the total content. The content of each is not less than 40.0% of the total.

Purified Stearic Acid USP is a mixture of the same acids which together constitute not less than 96.0% of the total content, and the content of $C_{18}H_{36}O_2$ is not less than 90.0% of the total.

Preparation—From edible fats and oils (see exception below) by boiling them with soda lye, separating the glycerin and decomposing the resulting soap with sulfuric or hydrochloric acid. The stearic acid subsequently is separated from any oleic acid by cold expression. It also is prepared by the hydrogenation and subsequent saponification of *olein*. It may be purified by recrystallization from alcohol.

Description—Hard, white or faintly yellowish somewhat glossy and crystalline solid, or a white or yellowish white powder; an odor and taste suggestive of tallow; melts about 55.5° and should not congeal at a temperature below 54°; the purified acid melts at 69 to 70° and congeals between 68 and 69°; slowly volatilizes between 90 and 100°.

Solubility—Practically insoluble in water; 1 g in about 20 mL of alcohol, 2 mL of chloroform, 3 mL of ether, 25 mL of acetone or 6 mL of carbon tetrachloride; freely soluble in carbon disulfide; also soluble in amyl acetate, benzene or toluene.

Incompatibilities—Insoluble stearates are formed with many *metals*. Ointment bases made with stearic acid may show evidence of drying out or lumpiness due to such a reaction when *zinc* or *calcium* salts are compounded therein.

Uses—In the preparation of sodium stearate which is the solidifying agent for the official glycerin suppositories, in enteric tablet coating, ointments and for many other commercial products, such as toilet creams, vanishing creams, solidified alcohol, etc. (When labeled solely for external use, it is exempt from the requirement that it be prepared from edible fats and oils.)

Other Emulsion Ointment Base Component

Wool Alcohols BP—Prepared by the saponification of the grease of the wool of sheep and separation of the fraction containing cholesterol and other alcohols. It contains not less than 30% cholesterol. Golden-brown solid, somewhat brittle when cold but becoming plastic when warm, with a faint characteristic odor; has a smooth and shiny fracture; melts not below 58°; acid value not more than 2; saponification value not

more than 12; emulsions made with this material do not darken on the surface or acquire an objectionable odor in hot weather. Insoluble in water; moderately soluble in alcohol; completely soluble in 25 parts of boiling anhydrous alcohol; freely soluble in ether, chloroform or petroleum ether. *Uses:* An emulsifying agent for the preparation of W/O emulsions; as a water absorbable substance in ointment bases; to improve the texture, stability and emollient properties of O/W emulsions. It is known also as *Lanolin Alcohols*.

Water-Soluble Ointment Bases and Components

Included in this section are bases prepared from the higher ethylene glycol polymers (PEG). These polymers are marketed under the trademark of Carbowax. The polymers have a wide range in molecular weight. Those with molecular weights ranging from 200 to 700 are liquids; those above 1000 are wax-like solids. The polymers are water-soluble, nonvolatile and unctuous agents. They do not hydrolyze or deteriorate and will not support mold growth. These properties account for their wide use in washable ointments. Mixtures of PEG are used to give bases of various consistency, such as very soft to hard bases for suppositories.

Glycol Ethers and Derivatives

This special class of ethers is of considerable importance in pharmaceutical technology. Both mono- and polyfunctional compounds are represented in the group. The simplest member is ethylene oxide [CH_2CH_2O], the internal or cyclic ether of the simplest glycol, ethylene glycol [$HOCH_2CH_2OH$]. External mono- and diethers of ethylene glycol [$ROCH_2CH_2OH$ and $ROCH_2CH_2OR'$] are well-known due largely to research done by the Carbide & Carbon.

Preparation—In the presence of NaOH at temperatures of the order of 120° to 135° and under a total pressure of about 4 atmospheres, ethylene oxide reacts with ethylene glycol to form compounds having the general formula $HOCH_2(CH_2OCH_2)_nCH_2OH$, commonly referred to as condensation polymers and termed polyethylene (or polyoxyethylene) glycols. Other glycols besides ethylene glycol function in similar capacity, and the commercial generic term adopted for the entire group is polyalkylene (or polyoxyalkylene) glycols.

Nomenclature—It is to be noted that these condensation polymers are bifunctional; ie, they contain both ether and alcohol linkages. The compound wherein $n = 1$ is the commercially important diethylene glycol [$HOCH_2CH_2OCH_2CH_2OH$], and its internal ether is the familiar dioxane [$CH_2CH_2OCH_2CH_2O$]. The mono- and diethers derived from diethylene glycol have the formulas $ROCH_2CH_2OCH_2CH_2OH$ and $ROCH_2CH_2OCH_2CH_2OR'$. The former commonly are termed "*Carbitols*" and the latter "*Cellosolves*," registered trademarks belonging to Carbide & Carbon Co.

Polyethylene glycols are differentiated in commercial nomenclature by adding a number to the name which represents the average molecular weight. Thus, polyethylene glycol 400 has an average molecular weight of about 400 (measured values for commercial samples range between 380 and 420) corresponding to a value of n for this particular polymer of approximately 8. Polymers have been produced in which the value of n runs into the hundreds. Up to n = approximately 15, the compounds are liquids at room temperature, viscosity and boiling point increasing with increasing molecular weight. Higher polymers are waxy solids and are termed commercially *Carbowaxes* (another Carbide & Carbon trademark).

It should be observed that the presence of the two terminal hydroxyl groups in the polyalkylene glycols makes possible the formation of both ether and ester derivatives, several of which are marketed products.

Uses—Because of their vapor pressure, solubility, solvent power, hygroscopicity, viscosity and lubricating characteristics, the polyalkylene glycols or their derivatives function in many applications as effective replacements for glycerin and water-insoluble oils. They find considerable use as plasticizers, lubricants, conditioners and finishing agents for processing textiles and rubber. They also are important as emulsifying agents and as dispersants for such diverse substances as dyes, oils, resins, insecticides and various types of pharmaceuticals. In addition, they are employed frequently as ingredients in ointment bases and in a variety of cosmetic preparations.

Polyethylene Glycols

Poly(oxy-1,2-ethanediyl), α-hydro ω-hydroxy-, Carbowaxes (*Carbide & Carbon*); Atpeg (*ICI*)

$$H\{OCH_2CH_2\}_nOH$$

Polyethylene glycols [25322-68-3].

Preparation—Ethylene glycol is reacted with ethylene oxide in the presence of NaOH at temperatures in the range of 120° to 135° under pressure of about 4 atm.

Description *Polyethylene glycols 200, 300, 400* and *600* are clear, viscous liquids at room temperature. *Polyethylene glycols 900, 1000, 1450, 3350, 4500* and *8000* are white, waxy solids. The glycols do not hydrolyze or deteriorate under typical conditions. As their molecular weight increases, their water solubility, vapor pressure, hygroscopicity and solubility in organic solvents decrease; at the same time, freezing or melting range, specific gravity, flash point and viscosity increase. If these compounds ignite, small fires should be extinguished with carbon dioxide or dry-chemical extinguishers and large fires with "alcohol"-type foam extinguishers.

Solubility—All members of this class dissolve in water to form clear solutions and are soluble in many organic solvents.

Uses—These possess a wide range of solubilities and compatibilities, which make them useful in pharmaceutical and cosmetic preparations. Their blandness renders them highly acceptable for hair dressings, hand lotions, sun-tan creams, leg lotions, shaving creams and skin creams (eg, a peroxide ointment which is stable may be prepared using these compounds, while oil-type bases inactivate the peroxide). Their use in washable ointments is discussed under *Ointments* (page 1602). They also are used in making suppositories, hormone creams, etc. See *Polyethylene Glycol Ointment* (below) and *Glycol Ethers* (above). The liquid polyethylene glycol 400 and the solid polyethylene glycol 3350, used in the proportion specified (or a permissible variation thereof) in the official Polyethylene Glycol Ointment, provide a water-soluble ointment base used in the formulation of many dermatological preparations. The solid, waxy, water-soluble glycols often are used to increase the viscosity of liquid polyethylene glycols and to stiffen ointment and suppository bases. In addition, they are used to compensate for the melting point-lowering effect of other agents, ie, chloral hydrate, etc, on such bases.

Polyethylene Glycol Ointment USP—*Preparation:* Heat polyethylene glycol 3350 (400 g) and polyethylene glycol 400 (600 g) on a water bath to 65°. Allow to cool, and stir until congealed. If a firmer preparation is desired, replace up to 100 g of polyethylene glycol 400 with an equal amount of polyethylene glycol 3350. If 6 to 25% of an aqueous solution is to be incorporated in this ointment, replace 50 g of polyethylene glycol 3350 by 50 g of stearyl alcohol. *Uses:* A water-soluble ointment base.

Polyoxyl 40 Stearate

Poly(oxy-1,2-ethanediyl), α-hydro-ω-hydroxy-, octadecanoate; Myrj (*ICI*)

$$RCOO(C_2H_4O)_nH$$

(*RCOO* is the stearate moiety, *n* is approximately 40)

Polyethylene glycol monostearate [9004-99-3]; a mixture of monostearate and distearate esters of mixed polyoxyethylene diols and corresponding free glycols, the average polymer length being equivalent to about 40 oxyethylene units. *Polyoxyethylene 50 Stearate* is a similar mixture in which the average polymer length is equivalent to about 50 oxyethylene units.

Preparation—One method consists of heating the corresponding polyethylene glycol with an equimolar portion of stearic acid.

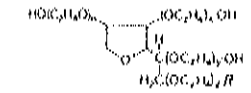
Description—White to light-tan waxy solid; odorless or has a faint fat-like odor; congeals between 37 and 47°.

Solubility—Soluble in water, alcohol, ether or acetone; insoluble in mineral or vegetable oils.

Uses—Contains ester and alcohol functions that impart both lyophilic and hydrophilic characteristics to make it useful as a surfactant and emulsifier. It is an ingredient of some water-soluble ointment and cream bases.

Polysorbates

Sorbitan esters, poly(oxy-1,2-ethanediyl) derivs. Manitans (*Ives-Cameron*); Sorlates (*Abbott*); Tweens (*ICI*)



[Sum of w, x, y, and z = 20, R = $(C_{17}H_{33})COO$.]

Sorbitan esters, polyoxyethylene derivatives; fatty acid esters of sorbitol and its anhydrides copolymerized with a varying number of moles of ethylene oxide. The NF recognizes: *Polysorbate 20* (*structure given above*), a laurate ester; *Polysorbate 40*, a palmitate ester; *Polysorbate 60*, a mixture of stearate and palmitate esters; and *Polysorbate 80*, an oleate ester.

Preparation—These important nonionic surfactants (page 268) are prepared starting with sorbitol by (1) elimination of water-forming sorbitan (a cyclic sorbitol anhydride); (2) partial esterification of the sorbitan with a fatty acid such as oleic or stearic acid yielding a hexitan ester known commercially as a *Span* and (3) chemical addition of ethylene oxide yielding a *Tween* (the polyoxyethylene derivative).

Description—*Polysorbate 80:* Lemon- to amber-colored, oily liquid; faint, characteristic odor; warm, somewhat bitter taste; specific gravity 1.07 to 1.09; pH (1:20 aqueous solution) 6 to 8.

Solubility—*Polysorbate 80:* Very soluble in water, producing an odorless and nearly colorless solution; soluble in alcohol, cottonseed oil, corn oil, ethyl acetate, methanol or toluene; insoluble in mineral oil.

Uses—Because of their hydrophilic and lyophilic characteristics, these nonionic surfactants are very useful as emulsifying agents forming O/W emulsions in pharmaceuticals, cosmetics and other types of products. Polysorbate 80 is an ingredient in *Coal Tar Ointment* and *Solution*. See *Glycol Ethers* (page 1313).

Other Water-Soluble Ointment Base Component

Polyethylene Glycol 400 Monostearate USP XVI—An ether, alcohol and ester. Semitransparent, whitish, odorless or nearly odorless mass; melts from 30 to 34°. Freely soluble in carbon tetrachloride, chloroform, ether or petroleum benzin; slightly soluble in alcohol; insoluble in water. *Uses:* A nonionic surface-active agent in the preparation of creams, lotions, ointments and similar pharmaceutical preparations, which are readily soluble in water.

Pharmaceutical Solvents

The remarkable growth of the solvent industry is attested by the more than 300 solvents now being produced on an industrial scale. Chemically, these include a great variety of organic compounds, ranging from hydrocarbons through alcohols, esters, ethers and acids to nitroparaffins. Their main applications are in industry and the synthesis of organic chemicals. Comparatively few, however, are used as solvents in pharmacy, because of their toxicity, volatility, instability and/or flammability. Those commonly used as pharmaceutical solvents are described in this section.

Acetone

2-Propanone; Dimethyl Ketone

CH_3COCH_3

Acetone [67-64-1] C_3H_6O (58.08).

Caution—It is very flammable. Do not use where it may be ignited.

Preparation—Formerly obtained exclusively from the destructive distillation of wood. The distillate, consisting principally of methanol, acetic acid and acetone was neutralized with lime and the acetone was separated from the methyl alcohol by fractional distillation. Additional quantities were obtained by pyrolysis of the calcium acetate formed in the neutralization of the distillate.

It now is obtained largely as a by-product of the butyl alcohol industry. This alcohol is formed in the fermentation of carbohydrates such as corn starch, molasses, etc, by the action of the bacterium *Clostridium acetobutylicum* (Weizmann fermentation) and it is always one of the products formed in the process. It also is obtained by the catalytic oxidation of isopropyl alcohol, which is prepared from propylene resulting from the "cracking" of crude petroleum.

Description—Transparent, colorless, mobile, volatile, flammable liquid with a characteristic odor; specific gravity not more than 0.789; distils between 55.5 and 57°; congeals about −95°; aqueous solution neutral to litmus.

Solubility—Miscible with water, alcohol, ether, chloroform or most volatile oils.

Uses—An *antiseptic* in concentrations above 80%. In combination with alcohol it is used as an antiseptic *cleansing* solution. It is employed as a menstruum in the preparation of oleoresins in place of ether. It is used as a *solvent* for dissolving fatty bodies, resins, pyroxylin, mercurials, etc, and also in the manufacture of many organic compounds such as chloroform, chlorobutanol and ascorbic acid.

Alcohol

Ethanol; Spiritus Vini Rectificatus; S. V. R.; Spirit of Wine; Methylcarbinol

Ethyl alcohol [64-17-5]; contains 92.3 to 93.8%, by weight (94.9 to 96.0%, by volume), at 15.56° (60°F) of C_2H_5OH (46.07).

Preparation—Has been made for centuries by fermentation of certain carbohydrates in the presence of *zymase*, an enzyme present in yeast cells. Usable carbohydrate-containing materials include molasses, sugar cane, fruit juices, corn, barley, wheat, potato, wood and waste sulfite liquors. As yeast is capable of fermenting only D-glucose, D-fructose, D-mannose and D-galactose it is essential that more complex carbohydrates, such as starch, be converted to one or more of these simple sugars before they can be fermented. This is accomplished variously, commonly by enzyme- or acid-catalyzed hydrolysis.

The net reaction that occurs when a hexose, glucose for example, is fermented to alcohol may be represented as

$$C_6H_{12}O_6 \rightarrow 2C_2H_5OH + 2CO_2$$

but the mechanism of the process is very complex. The fermented liquid, containing about 15% of alcohol, is distilled to obtain a distillate containing 94.9% of C_2H_5OH, by volume. To produce *absolute alcohol*, the 95% product is dehydrated by various processes.

It may be produced also by hydration of ethylene, abundant supplies of which are available from natural and coke oven gases, from waste gases of the petroleum industry and other sources. In another synthesis acetylene is hydrated catalytically to acetaldehyde, which then is hydrogenated catalytically to ethyl alcohol.

Description—Transparent, colorless, mobile, volatile liquid; slight but characteristic odor; burning taste; boils at 78° but volatilizes even at a low temperature, and is flammable; when pure, it is neutral towards all indicators; specific gravity at 15.56° (the US Government standard temperature for Alcohol) not above 0.816, indicating not less than 92.3% of C_2H_5OH by weight or 94.9% by volume.

Solubility—Miscible with water, acetone, chloroform, ether or many other organic solvents.

Incompatibilities—Thin and preparations containing a high percentage of alcohol will precipitate many inorganic salts from an aqueous solution. *Acacia* generally is precipitated from a hydroalcoholic medium when the alcohol content is greater than about 35%.

Strong *oxidizing agents* such as *chlorine, nitric acid, permanganate* or *chromate* in acid solution react, in some cases violently, with it to produce oxidation products.

Alkalies cause a darkening in color due to the small amount of aldehyde usually present in it.

Uses—In pharmacy principally for its solvent powers (page 216). It also is used as the starting point in the manufacture of many important compounds, like ether, chloroform, etc. It also is used as a fuel, chiefly in the denatured form.

It is a CNS depressant. Consequently, it occasionally has been administered intravenously for preoperative and postoperative sedation in patients in whom other measures are ineffective or contraindicated. The dose employed is 1 to 1.5 mL/kg. Its intravenous use is a specialized procedure and should be employed only by one experienced in the technique of such use.

It is used widely and abused by lay persons as a sedative. It has, however, no medically approved use for this purpose. Moreover, alcohol potentiates the CNS effects of numerous sedative and depressant drugs. Hence, it should not be used by patients taking certain prescription drugs or OTC medications (see page 1862).

Externally, it has a number of medical uses. It is a solvent for the toxicodendrol causing *ivy poisoning*, and should be used to wash the skin thoroughly soon after contact. In a concentration of 25% it is employed for bathing the skin for the purpose of *cooling* and *reducing fevers*. In high concentrations it is a *rubefacient* and an ingredient of many liniments. In a concentration of 50% it is used to prevent sweating in *astringent* and *antihidrotic* lotions. It also is employed to cleanse and harden the skin and is helpful in preventing *bedsores* in bedridden patients. In a concentration of 60 to 90% it is germicidal. At optimum concentration (70% by weight) it is a good *antiseptic* for the skin (*local anti-infective*) and also for instruments. It also is used as a *solvent* to cleanse the skin splashed with phenol. High concentrations of it often are injected into nerves and ganglia for the *relief of pain*, accomplishing this by causing nerve degeneration.

Denatured Alcohol

An act of Congress June 7, 1906, authorizes the withdrawal of alcohol from bond without the payment of internal revenue tax, for the purpose of denaturation and use in the arts and industries. This is ethyl alcohol to which have been added such denaturing materials as to render the alcohol unfit for use as an intoxicating beverage. It is divided into two classes, namely, *completely denatured alcohol* and *specially denatured alcohol*, prepared in accordance with approved formulas prescribed in Federal Industrial Alcohol Regulations 3.

Information regarding the use of alcohol and permit requirements may be obtained from the Regional Director, Bureau of Alcohol, Tobacco and Firearms, in any of the following offices: Cincinnati, OH; Philadelphia, PA; Chicago, IL; New York, NY; Atlanta, GA; Dallas, TX and San Francisco, CA. Federal regulation provides that completely and specially denatured alcohols may be purchased by properly qualified persons from duly established denaturing plants or bonded dealers. No permit is required for the purchase and use of completely denatured alcohol unless the purchaser intends to recover the alcohol.

Completely Denatured Alcohol—This term applies to ethyl alcohol to which has been added materials (methyl isobutyl ketone, pyronate, gasoline, acetaldol, kerosene, etc) of such nature that the products may be sold and used within certain limitations without permit and bond.

Specially Denatured Alcohol—This alcohol is intended for use in a greater number of specified arts and industries than completely denatured alcohol and the character of the denaturant or denaturants used is such that specially denatured alcohol may be sold, possessed and used only by those persons or firms that hold basic permits and are covered by bond.

Formulas for products using specially denatured alcohol must be approved prior to use by the Regional Director, Bureau of Alcohol, Tobacco and Firearms in any of the regional offices listed above.

Uses—Approximately 50 specially denatured alcohol formulas containing combinations of more than 90 different denaturants are available to fill the needs of qualified users. Large amounts of specially denaturated alcohols are used as raw materials in the production of acetaldehyde, synthetic rubber, vinegar and ethyl chloride as well as in the manufacture of proprietary solvents and cleaning solutions. Ether and chloroform can be made from suitably denatured alcohols and formulas for the manufacture of Iodine Tincture, Green Soap Tincture and Rubbing Alcohol are set forth in the regulations.

Specially denatured alcohols also are used as solvents for surface coatings, plastics, inks, toilet preparations and external pharmaceuticals. Large quantities are used in the processing of such food and drug products as pectin, vitamins, hormones, antibiotics, alkaloids and blood products. Other uses include supplemental motor fuel, rocket and jet fuel, antifreeze solutions, refrigerants and cutting oils. Few products are manufactured today that do not require the use of alcohol at some stage of production. Specially denatured alcohol may not be used in the manufacture of foods or internal medicines where any of the alcohol remains in the finished product.

Rose Water Ointment

Cold Cream; Galen's Cerate

Cetyl Esters Wax	125 g
White Wax	120 g
Almond Oil	560 g
Sodium Borate	5 g
Stronger Rose Water	25 mL
Purified Water	165 mL
Rose Oil	0.2 mL
To make about	1000 g

Reduce the cetyl esters wax and the white wax to small pieces, melt them on a steam bath, add the almond oil and continue heating until the temperature of the mixture reaches 70°. Dissolve the sodium borate in the purified water and stronger rose water, warmed to 70°, and gradually add the warm solution to the melted mixture, stirring rapidly and continuously until it has cooled to about 45°. Incorporate the rose oil.

It must be free from rancidity. If the ointment has been chilled, warm it slightly before attempting to incorporate other ingredients (see USP for allowable variations).

History—Originated by Galen, the famous Roman physician-pharmacist of the 1st century AD, was known for many centuries by the name of *Unguentum or Ceratum Refrigerans*. It has changed but little in proportions or method of preparation throughout many centuries.

Uses—An *emollient* and *ointment base*. It is a W/O emulsion.

Diluted Alcohol

Diluted Ethanol

A mixture of alcohol and water containing 41.0 to 42.0%, by weight (48.4 to 49.5%, by volume), at 15.56°, of C_2H_5OH (46.07).

Preparation—

Alcohol	500 mL
Purified Water	500 mL

Measure the alcohol and the purified water separately at the same temperature, and mix. If the water and the alcohol and the resulting mixture are measured at 25°, the volume of the mixture will be about 970 mL.

When equal volumes of alcohol and water are mixed together, a rise in temperature and a contraction of about 3% in volume take place. In small operations the contraction generally is disregarded; in larger operations it is very important. If 50 gal of official alcohol are mixed with 50 gal of water, the product will not be 100 gal of diluted alcohol, but only 96⅛ gal, a contraction of 3¾ gal. *USP Proof Spirit* differs from this and is stronger; it contains 50%, by volume, of absolute alcohol at 15.56° (60°F). This corresponds to 42.5% by weight, and has a specific gravity of 0.9341 at the same temperature. If spirits have a specific gravity lower than that of "proof spirit" (0.9341), they are said to be "*above proof*," if greater, "*below proof*."

It also may be prepared from the following:

Alcohol	408 g
Purified Water	500 g

Rules for Dilution—The following rules are applied when making an alcohol of any required lower percentage from an alcohol of any given higher percentage:

I. By Volume—Designate the volume percentage of the stronger alcohol by *V*, and that of the weaker alcohol by *v*.

Rule—Mix *v* volumes of the stronger alcohol with purified water to make *V* volumes of product. Allow the mixture to stand until full contraction has taken place, and until it has cooled, then make up the deficiency in the *V* volumes by adding more purified water.

Example—An alcohol of 30% by volume is to be made from an alcohol of 94.9% by volume.—Take 30 volumes of the 94.9% alcohol, and add enough purified water to produce 94.9 volumes at room temperature.

II. By Weight—Designate the weight-percentage of the stronger alcohol by *W*, and that of the weaker alcohol by *w*.

Rule—Mix *w* parts by weight of the stronger alcohol with purified water to make *W* parts by weight of product.

Example—An alcohol of 50% by weight is to be made from an alcohol of 92.3% by weight.—Take 50 parts by weight of the 92.3% alcohol, and add enough purified water to produce 92.3 parts by weight.

Description—As for *Alcohol*, except its specific gravity is 0.935 to 0.937 at 15.56°, indicating that the strength of C_2H_5OH corresponds to that given in the official definition.

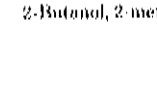
Uses—A menstruum in making tinctures, fluidextracts, extracts, etc. Its properties already have been described fully in connection with the various preparations. Its value consists not only in its *antiseptic* properties, but also in its possessing the *solvent* powers of both water and alcohol. See *Alcohol*.

Nonbeverage Alcohol

This is tax-paid alcohol or distilled spirits used in the manufacture, by approved formula, of such medicines, medicinal preparations, food products, flavors or flavoring extracts as are unfit for beverage purposes. Internal Revenue Service Regulations provide that qualified holders of Special Tax Stamps who use tax paid alcohol or distilled spirits in the types of products listed above, may file a claim for *alcohol tax drawback* or refund of a considerable part of the tax paid.

Amylene Hydrate

2-Butanol, 2-methyl-, Tertiary Amyl Alcohol; Dimethylethylcarbinol



tert-Pentyl alcohol [75-85-4] $C_5H_{12}O$ (88.15).

Preparation—Amylene is mixed with 2 volumes of 60% H_2SO_4, both previously cooled to 0°, for about 1 hr; then neutralized with soda, distilled and the first half of the distillate containing most of the amylene hydrate is treated with anhydrous potassium carbonate and redistilled.

Description—Clear, colorless liquid of camphoraceous odor; solution neutral to litmus; specific gravity 0.803 to 0.807; distils completely between 97 and 103°.

Solubility—1 g in about 8 mL of water; miscible with alcohol, chloroform, ether or glycerin.

Uses—Chiefly, a *pharmaceutic necessity for Tribromoethanol Solution* (RPS-15, page 985). It has been used as a *sedative-hypnotic* in doses of 1 to 4 g administered in glycerin.

Chloroform—page 1320.
Ether—page 1041.
Ethyl Acetate—page 1294.

Glycerin

1,2,3-Propanetriol; Glycerol

Glycerol [56-81-5] $C_3H_8O_3$ (92.09).

Chemically, it is the simplest trihydric alcohol. It is worthy of special note because the two terminal alcohol groups are primary, whereas the middle one is secondary. Thus this becomes the first polyhydric alcohol which can yield both an aldose (*glyceraldehyde*) and a ketose (*dihydroxyacetone*).

Preparation—

1. By saponification of fats and oils in the manufacture of soap.
2. By hydrolysis of fats and oils through pressure and superheated steam.
3. By fermentation of beet sugar molasses in the presence of large amounts of sodium sulfite. Under these conditions a reaction takes place expressed as

$$C_6H_{12}O_6 \rightarrow C_3H_5(OH)_3 + CH_3CHO + CO_2$$
Glucose — Glycerin — Acetaldehyde

4. Glycerin is now prepared in large quantities from propylene, a petroleum product. This hydrocarbon is chlorinated at about 400° to form allyl chloride, which is converted to allyl alcohol. Treatment of the unsaturated alcohol with hypochlorous acid [HOCl] yields the chlorohydrin derivative. Extraction of HCl with soda lime yields 2,3-epoxypropanol which undergoes hydration to glycerin.

Description—Clear, colorless, syrupy liquid with a sweet taste and not more than a slight, characteristic odor, which is neither harsh nor disagreeable; when exposed to moist air it absorbs water and also such gases as H_2S and SO_2; solutions are neutral; specific gravity not below 1.249 (not less than 95% $C_3H_5(OH)_3$); boils at about 290° under 1 atm, with decomposition, but can be distilled intact in a vacuum.

Solubility—Miscible with water, alcohol methanol; 1 g in about 12 mL of ethyl acetate or about 15 mL of acetone; insoluble in chloroform, ether or fixed and volatile oils.

Incompatibilities—An explosion may occur if it is triturated with strong *oxidizing agents* such as *chromium trioxide*, *potassium chlorate* or *potassium permanganate*. In dilute solutions the reactions proceed at a slower rate forming several oxidation products. Iron is an occasional contaminant of it and may be the cause of a darkening in color in mixtures containing *phenols*, *salicylates*, *tannin*, etc.

With *boric acid* or *sodium borate*, it forms a complex, generally spoken of as glyceroboric acid, which is a much stronger acid than boric acid.

Uses—One of the most valuable products known to pharmacy by virtue of its *solvent* property. It is useful as a *humectant* in keeping substances moist, owing to its hygroscopicity. Its agreeable taste and high viscosity adapt it for many purposes. Some modern ice collars and ice bags contain it and water hermetically sealed within vulcanized rubber bags. The latter are sterilized by dipping in a germicidal solution and are stored in the refrigerator until needed. It also has some therapeutic uses. In pure anhydrous form, it is used in the eye to reduce corneal edema and to facilitate ophthalmoscopic examination. It is used orally as an evacuant and, in 50 to 75% solution, as a systemic osmotic agent.

Isopropyl Alcohol—page 1167.

Methyl Alcohol

Methanol; Wood Alcohol

CH_3OH

Methanol [67-56-1] CH_4O (32.04).

Caution—It is poisonous.

Preparation—By the catalytic reduction of carbon monoxide or carbon dioxide with hydrogen. A zinc oxide-chromium oxide catalyst is used commonly.

Description—Clear, colorless liquid; characteristic odor; flammable; specific gravity not more than 0.796; distils within a range of 1° between 63.5 and 65.7°.

Solubility—Miscible with water, alcohol, ether, benzene or most other organic solvents.

Uses—*A pharmaceutic aid* (solvent). It is toxic. Ingestion may result in blindness; vapors also may cause toxic reactions.

Methyl Isobutyl Ketone

2-Pentanone, 4-methyl-,

$(CH_3)_2CHCH_2COCH_3$

4-Methyl-2-pentanone [108-10-1]; contains not less than 99% of $C_6H_{12}O$ (100.16).

Description—Transparent, colorless, mobile, volatile liquid; faint, ketonic and camphoraceous odor, distils between 114 and 117°.

Solubility—Slightly soluble in water; miscible with alcohol, ether or benzene.

Uses—A *denaturant* for rubbing alcohol and also a *solvent* for gums, resins, nitrocellulose, etc. It may be irritating to the eyes and mucous membranes, and, in high concentrations, narcotic.

Monoethanolamine

Ethanol, 2-amino-, Ethanolamine; Ethylolamine

$HOCH_2CH_2NH_2$

2-Aminoethanol [141-43-5] C_2H_7NO (61.08).

Preparation—This alkanolamine is prepared conveniently by treating ethylene oxide with ammonia.

Description—Clear, colorless, moderately viscous liquid; distinctly ammoniacal odor; affected by light; specific gravity 1.013 to 1.016; distils between 167 and 173°.

Solubility—Miscible in all proportions with water, acetone, alcohol, glycerin or chloroform; immiscible with ether, solvent hexane or fixed oils; dissolves many essential oils.

Uses—A *solvent* for fats, oils and many other substances, it is a pharmaceutical necessity for *Thimerosal Solution* (page 1173). It combines with fatty acids to form soaps which find application in various types of emulsions such as lotions, creams, etc.

Propylene Glycol

$CH_3CH(OH)CH_2OH$

1,2-Propanediol [57-55-6] $C_3H_8O_2$ (76.10).

Preparation—Propylene is converted successively to its chlorohydrin (with HOCl), epoxide (with Na_2CO_3) and glycol (with water in presence of protons).

Description—Clear, colorless, viscous and practically odorless liquid; slightly acrid taste; specific gravity 1.035 to 1.037; completely distils between 184 and 189°; absorbs moisture from moist air.

Solubility—Miscible with water, alcohol, acetone or chloroform; soluble in ether; dissolves many volatile oils; immiscible with fixed oils.

Uses—A *solvent*, *preservative* and *humectant*. See *Hydrophilic Ointment* (page 1311).

Trolamine

Ethanol, 2,2',2''-nitrilotris-, Triethanolamine

2,2',2''-Nitrilotriethanol [102-71-6] $N(C_2H_4OH)_3$ (149.19); a mixture of alkanolamines consisting largely of triethanolamine, containing some diethanolamine $[NH(C_2H_4OH)_2 = 105.14]$ and monoethanolamine $[NH_2C_2H_4OH = 61.08]$.

Preparation—Along with some mono- and diethanolamine, by the action of ammonia on ethylene oxide.

Description—Colorless to pale yellow, viscous, hygroscopic liquid; slight odor of ammonia; aqueous solution is very alkaline; melts about 21°; specific gravity 1.120 to 1.128; a strong base and readily combines even with weak acids to form salts.

Solubility—Miscible with water or alcohol; soluble in chloroform; slightly soluble in ether or benzene.

Uses—In combination with a fatty acid, eg, oleic acid (see *Benzyl Benzoate Lotion*, page 1246), as an *emulsifier*. See *Monoethanolamine*.

Water—page 1300.

Other Pharmaceutical Solvents

Alcohol, Dehydrated, BP, PhI [Dehydrated Ethanol; Absolute Alcohol]—Transparent, colorless, mobile, volatile liquid; characteristic odor; burning taste; specific gravity not more than 0.798 at 15.56°; hygroscopic, flammable and boils about 78°C. Miscible with water, ether or chloroform. *Uses:* A pharmaceutical solvent; also used by injection for relief of pain (see *Alcohol*, page 1314).

Coconut Oil [Coconut Oil; Copra Oil]—The fixed oil obtained by expression or extraction from the kernels of the seeds of *Cocos nucifera* Linné (Fam *Palmae*). Pale yellow to colorless liquid between 28 and 30°, semisolid at 20° and a hard, brittle crystalline solid below 15°; odorless and tasteless or has a faint odor and taste characteristic of coconut; it must not be used if it has become rancid; melts about 23°; specific gravity 0.918 to 0.923. Readily soluble in alcohol, ether, chloroform, carbon disulfide or petroleum benzin; insoluble in water.

Petroleum Benzin [Petroleum ether; Purified benzin]—Clear, colorless, volatile liquid; ethereal or faint, petroleum-like odor; neutral reaction; specific gravity 0.634 to 0.660. Practically insoluble in water; miscible with ether, chloroform, benzene or fixed oils. *Caution:* Highly flammable, and its vapor, when mixed with air and ignited, may explode. *Uses:* A solvent for fats, resins, oils and similar substances.

Miscellaneous Pharmaceutical Necessities

The agents listed in this section comprise a heterogeneous group of substances with both pharmaceutical and industrial applications. Pharmaceutically, some of these agents are used as diluents, enteric coatings, excipients, filtering agents and as ingredients in products considered in other chapters. Industrially, some of these agents are used in various chemical processes, in the synthesis of other chemicals and in the manufacture of fertilizers, explosives, etc.

Acetic Acid

Acetic acid; a solution containing 36 to 37%, by weight, of $C_2H_4O_2$ (60.05).

Preparation—By diluting with distilled water an acid of higher concentration, such as the 80% product, or more commonly glacial acetic acid, using 350 mL of the latter for the preparation of each 1000 mL of acetic acid.

Description—Clear, colorless liquid, having a strong characteristic odor and a sharply acid taste; specific gravity about 1.045; congeals about −14°; acid to litmus.

Solubility—Miscible with water, alcohol or glycerin.

Uses—In pharmacy as a *solvent* and *menstruum*, and for making diluted acetic acid. It also is used as a starting point in the manufacture of many other organic compounds, eg, acetates, acetanilid, sulfonamides, etc. It is official primarily as a *pharmaceutic necessity* for the preparation of *Aluminum Subacetate Solution* (RPS-17, page 778).

Diluted Acetic Acid

Dilute Acetic Acid

A solution containing, in each 100 mL, 5.7 to 6.3 g of $C_2H_4O_2$.

Preparation—

Acetic Acid	158 mL
Purified Water, a sufficient quantity,	
To make	1000 mL

Mix the ingredients.

Note—This acid also may be prepared by diluting 58 mL of glacial acetic acid with sufficient purified water to make 1000 mL.

Description—Essentially the same properties, solubility, purity and identification reactions as *Acetic Acid*, but its specific gravity is about 1.008 and it congeals about −2°.

Uses—*Bactericidal* to many types of microorganisms and occasionally is used in 1% solution for surgical dressings of the skin. A 1% solution is *spermatocidal*. It also is used in vaginal douches for the management of *Trichomonas*, *Candida* and *Hemophilus* infections.

Glacial Acetic Acid

Concentrated Acetic Acid; Crystallizable Acetic Acid; Ethanoic Acid; Vinegar Acid

CH_3COOH

Glacial acetic acid [64-19-7] $C_2H_4O_2$ (60.05).

Preparation—This acid is termed "glacial" because of its solid, glassy appearance when congealed. In one process it is produced by distillation of weaker acids to which has been added a water-entraining substance such as ethylene dichloride. In this method, referred to as "azeotropic distillation," the ethylene dichloride dis-

tils out with the water before the acid distils over, thereby effecting concentration of the latter.

In another process the aqueous acid is mixed with triethanolamine and heated. The acid combines with the triethanolamine to form a triethanolamine acetate. The water is driven off first; then, at a higher temperature, the triethanolamine compound decomposes to yield this acid.

A greater part of the acid now available is made synthetically from acetylene. When acetylene is passed into this acid containing a metallic catalyst such as mercuric oxide, ethylidene diacetate is produced which yields, upon heating, acetic anhydride and acetaldehyde. Hydration of the former and air oxidation of the latter yield this acid.

Description—Clear, colorless liquid; pungent, characteristic odor; when well-diluted with water, it has an acid taste; boils about 118°; congeals at a temperature not lower than 15.6°, corresponding to a minimum of 99.4% of CH_3COOH; specific gravity about 1.05.

Solubility—Miscible with water, alcohol, acetone, ether or glycerin; insoluble in carbon tetrachloride or chloroform.

Uses—A *caustic* and *vesicant* when applied externally and is often sold under various disguises as a *corn solvent*. It is an excellent *solvent* for fixed and volatile oils and many other organic compounds. It is used primarily as an *acidifying agent*.

Almond Oil—RPS-16, page 720.

Aluminum

Aluminum Al (26.98); the free metal in the form of finely divided powder. It may contain oleic acid or stearic acid as a lubricant. It contains not less than 95% of Al, and not more than 5% of *Acid-insoluble substances*, including any added fatty acid.

Description—Very fine, free-flowing, silvery powder free from gritty or discolored particles.

Solubility—Insoluble in water or alcohol; soluble in hydrochloric and sulfuric acids or in solutions of fixed alkali hydroxides.

Uses—A *protective*. An ingredient in *Aluminum Paste* (RPS-14, page 772).

Aluminum Monostearate

Aluminum, dihydroxy(octadecanoato-*O*)-,

Dihydroxy(stearato)aluminum [7047-84-9]; a compound of aluminum with a mixture of solid organic acids obtained from fats, and consists chiefly of variable proportions of aluminum monostearate and aluminum monopalmitate. It contains the equivalent of 14.5 to 16.5% of Al_2O_3 (101.96).

Preparation—By interaction of a hydroalcoholic solution of potassium stearate with an aqueous solution of potassium alum, the precipitate being purified to remove free stearic acid and some aluminum distearate simultaneously produced.

Description—Fine, white to yellowish white, bulky powder; faint, characteristic odor.

Solubility—Insoluble in water, alcohol or ether.

Uses—A *pharmaceutical necessity* used in the preparation of *Sterile Procaine Penicillin G with Aluminum Stearate Suspension* (see page 1191).

Strong Ammonia Solution

Stronger Ammonia Water; Stronger Ammonium Hydroxide Solution; Spirit of Hartshorn

Ammonia [1336-21-6]; a solution of NH_3 (17.03), containing 27.0 to 31.0% (*w/w*) of NH_3. Upon exposure to air it loses ammonia rapidly.

Caution—Use care in handling it because of the caustic nature of the Solution and the irritating properties of its vapor. Cool the container well before opening, and cover the closure with a cloth or similar material while opening. Do not taste it, and avoid inhalation of its vapor.

Preparation—Ammonia is obtained commercially chiefly by synthesis from its constituent elements, nitrogen and hydrogen, combined under high pressure and at high temperature in the presence of a catalyst.

Description—Colorless, transparent liquid; exceedingly pungent, characteristic odor; even when well-diluted it is strongly alkaline to litmus; specific gravity about 0.90.

Solubility—Miscible with alcohol.

Uses—Only for chemical and pharmaceutical purposes. It is used primarily in making ammonia water by dilution and as a chemical reagent. It is too strong for internal administration. It is an ingredient in *Aromatic Ammonia Spirit* (page 1533).

Bismuth Subnitrate

Basic Bismuth Nitrate; Bismuth Oxynitrate; Spanish White; Bismuth Paint; Bismuthyl Nitrate

Bismuth hydroxide nitrate oxide [1304-85-4] $Bi_6O(OH)_9(NO_3)_5$ (1461.99); a basic salt which, dried at 105° for 2 hr, yields upon ignition not less than 79% of Bi_2O_3 (465.96).

Preparation—A solution of bismuth nitrate is added to boiling water to produce the subnitrate by hydrolysis.

Description—White, slightly hygroscopic powder; suspension in distilled water is faintly acid to litmus (pH about 5).

Solubility—Practically insoluble in water or organic solvents; dissolves readily in an excess of hydrochloric or nitric acid.

Incompatibilities—Slowly hydrolyzed in *water* with liberation of nitric acid; thus, it possesses the incompatibilities of the acid. *Reducing agents* darken it with the production of metallic bismuth.

Uses—A *pharmaceutical necessity* in the preparation of milk of bismuth. It also is used as an *astringent*, *adsorbent* and *protective*; however, its value as a protective is questionable. This agent, like other insoluble bismuth salts, is used topically in lotions and ointments.

Barium Hydroxide Lime

A mixture of barium hydroxide octahydrate and calcium hydroxide. It also may contain potassium hydroxide and may contain an indicator that is inert toward anesthetic gases such as ether, cyclopropane and nitrous oxide, and that changes color when the barium hydroxide lime no longer can absorb carbon dioxide.

Caution—Since it contains a soluble form of barium, it is toxic if swallowed.

Description—White or grayish white granules; may have a color if an indicator has been added.

Uses—A carbon dioxide adsorbent. See *Soda Lime* (page 1325).

Boric Acid

Boric acid (H_3BO_3); Boracic Acid; Orthoboric Acid

Boric acid [10043-35-3] H_3BO_3 (61.83).

Preparation—Lagoons of the volcanic districts of Tuscany formerly furnished the greater part of this acid and borax of commerce. Borax is now found native in California and some of the other western states; calcium and magnesium borates are found there also. It is produced from native borax, or from the other borates, by reacting with hydrochloric or sulfuric acid.

Description—Colorless scales of a somewhat pearly luster, or crystals, but more commonly a white powder slightly unctuous to the touch; odorless and stable in the air; volatilizes with steam.

Solubility—1 g in 18 mL of water, 18 mL of alcohol, 4 mL of glycerin, 4 mL of boiling water or 6 mL of boiling alcohol.

Uses—A buffer, and it is this use that is recognized officially. It is a very weak *germicide* (*local anti-infective*). Its nonirritating properties make its solutions suitable for application to such delicate structures as the cornea of the eye. Aqueous solutions are employed as an eye wash, mouth wash and for irrigation of the bladder. A 2.2% solution is isotonic with lacrimal fluid. Solutions, even if they are made isotonic, will hemolyze red blood cells. It also is employed as a dusting powder, when diluted with some inert material. It can be absorbed through irritated skin, eg, infants with diaper rash.

Although it is not absorbed significantly from intact skin, it is absorbed from damaged skin and fatal poisoning, particularly in infants, has occurred with topical application to burns, denuded areas, granulation tissue and serous cavities. *Serious poisoning can*

result from oral ingestion of as little as 5 g. Symptoms of poisoning are nausea, vomiting, abdominal pain, diarrhea, headache and visual disturbance. Toxic alopecia has been reported from the chronic ingestion of a mouth wash containing it. The kidney may be injured and death may result. Its use as a preservative in beverages and foods is prohibited by national and state legislation. *There is always present the danger of confusing it with dextrose when compounding milk formulas for infants. Fatal accidents have occurred.* For this reason boric acid in bulk is colored, so that it cannot be confused with dextrose.

It is used to prevent discoloration of physostigmine solutions.

Dose—*Topically*, as required.

Calcium Hydroxide

Slaked Lime; Calcium Hydrate

Calcium hydroxide [1305-62-0] $Ca(OH)_2$ (74.09).

Preparation—By reacting freshly prepared calcium oxide with water.

Description—White powder; alkaline, slightly bitter taste; absorbs carbon dioxide from the air forming calcium carbonate; solutions exhibit a strong alkaline reaction.

Solubility—1 g in 630 mL of water or 1300 mL of boiling water; soluble in glycerin or syrup; insoluble in alcohol; the solubility in water is decreased by the presence of fixed alkali hydroxides.

Uses—In the preparation of *Calcium Hydroxide Solution*.

Calcium Hydroxide Topical Solution

Calcium Hydroxide Solution; Lime Water

A solution containing, in each 100 mL, not less than 140 mg of $Ca(OH)_2$ (74.09).

Note—The solubility of calcium hydroxide varies with the temperature at which the solution is stored, being about 170 mg/100 mL at 15°, and less at a higher temperature. The official concentration is based upon a temperature of 25°.

Preparation—

Calcium Hydroxide	3 g
Purified Water	1000 mL

Add the calcium hydroxide to 1000 mL of cool, purified water, and agitate the mixture vigorously and repeatedly during 1 hr. Allow the excess of calcium hydroxide to settle. Dispense only the clear, supernatant liquid.

The undissolved portion of the mixture is not suitable for preparing additional quantities of the solution.

The object of keeping lime water over undissolved calcium hydroxide is to insure a saturated solution.

Description—Clear, colorless liquid; alkaline taste; strong alkaline reaction; absorbs carbon dioxide from the air, a film of calcium carbonate forming on the surface of the liquid; when heated, it becomes turbid, owing to the separation of calcium hydroxide, which is less soluble in hot than in cold water.

Uses—Too dilute to be effective as a gastric antacid. It is employed *topically* as a *protective* in various types of lotions. In some lotion formulations it is used with olive oil or oleic acid to form calcium oleate that functions as an emulsifying agent. The USP classes it as an *astringent*.

Dose—*Topically*, in astringent solutions and lotions as required (see *Calamine Lotion*, page 762).

Calcium Pantothenate, Racemic—page 1022.

Calcium Stearate

Octadecanoic acid, calcium salt

Calcium stearate [1592-23-0]; a compound of calcium with a mixture of solid organic acids obtained from fats and consists chiefly of variable proportions of stearic and palmitic acids [calcium stearate, $C_{36}H_{70}CaO_4$ = 607.03; calcium palmitate, $C_{32}H_{62}CaO_4$ = 550.92]; contains the equivalent of 9 to 10.5% of CaO (calcium oxide).

Preparation—By precipitation from interaction of solutions of calcium chloride and the sodium salts of the mixed fatty acids (stearic and palmitic).

Description—Fine, white to yellowish white, bulky powder; slight, characteristic odor; unctuous and free from grittiness.

Solubility—Insoluble in water, alcohol or ether.

Uses—A *lubricant* in the manufacture of compressed tablets. It also is used as a conditioning agent in food and pharmaceutical products. Its virtually nontoxic nature and unctuous properties makes it ideal for these purposes.

Calcium Sulfate

Sulfuric acid, calcium salt (1:1); Gypsum; Terra Alba

Calcium sulfate (1:1) [7778-18-9] $CaSO_4$ (136.14); *dihydrate* [10101-41-4] (172.17).

Preparation—From natural sources or by precipitation from interaction of solutions of calcium chloride and a soluble sulfate.

Description—Fine, white to slightly yellow-white, odorless powder.

Solubility—Dissolves in dilute HCl; slightly soluble in water.

Uses—A *diluent* in the manufacture of compressed tablets. It is sufficiently inert that few undesirable reactions occur in tablets made with this substance. It also is used for making plaster casts and supports.

Carnauba Wax

Obtained from the leaves of *Copernicia cerifera* Mart (Fam *Palmae*).

Preparation—Consists chiefly of *myricyl cerotate* with smaller quantities of *myricyl alcohol*, *ceryl alcohol* and *cerotic acid*. It is obtained by treating the leaf buds and leaves of *Copernicia cerifera*, the so-called *Brazilian Wax Palm*, with hot water.

Description—Light-brown to pale yellow, moderately coarse powder; characteristic bland odor; free from rancidity; specific gravity about 0.99; melts about 84°.

Solubility—Insoluble in water; freely soluble in warm benzene; soluble in warm chloroform or toluene; slightly soluble in boiling alcohol.

Uses—A pharmaceutic aid used as a *polishing agent* in the manufacture of coated tablets.

Microcrystalline Cellulose

Cellulose [9004-34-6]; purified, partially depolymerized cellulose prepared by treating alpha cellulose, obtained as a pulp from fibrous plant material, with mineral acids.

Preparation—Cellulose is subjected to the hydrolytic action of 2.5 *N* HCl at the boiling temperature of about 105° for 15 min, whereby amorphous cellulosic material is removed and aggregates of crystalline cellulose are formed. These are collected by filtration, washed with water and aqueous ammonia and disintegrated into small fragments, often termed cellulose crystallites, by vigorous mechanical means such as a blender. US Pat. 3,141,875.

Description—Fine, white, odorless, crystalline powder; consists of free-flowing, nonfibrous particles.

Solubility—Insoluble in water, dilute acids or most organic solvents; slightly soluble in NaOH solution (1 in 20).

Uses—A tablet diluent and disintegrant. It can be compressed into self-binding tablets which disintegrate rapidly when placed in water.

Microcrystalline Cellulose and Sodium Carboxymethylcellulose—A colloid-forming, attrited mixture of microcrystalline cellulose and sodium carboxymethylcellulose. Tasteless, odorless, white to off-white, coarse to fine powder; pH (dispersion) 6 to 8. Swells in water, producing, when dispersed, a white, opaque dispersion or gel; insoluble in organic solvents or dilute acids. *Uses*: Pharmaceutic aid (suspending agent). *Grades Available* (amounts of sodium carboxymethylcellulose producing viscosities in the concentrations designated): 8.5%, 120 cps in 2.1% solution; 11%, 120 cps in 1.2% solution; 11%, 65 cps in 1.2% solution.

Powdered Cellulose—page 1305.

Cellulose Acetate Phthalate

Cellulose, acetate, 1,2-benzenedicarboxylate

Cellulose acetate phthalate [9004-38-0]; a reaction product of the phthalic anhydride and a partial acetate ester of cellulose. When

dried at 105° for 2 hr, it contains 19 to 23.5% of acetyl (C_2H_3O) groups and 30 to 36.0% of phthalyl (*o*-carboxybenzoyl, $C_8H_5O_3$) groups.

Preparation—Cellulose is esterified by treatment with acetic and phthalic acid anhydrides.

Description—Free-flowing, white powder; may have a slight odor of acetic acid.

Solubility—Insoluble in water or alcohol; soluble in acetone or dioxane.

Uses—An *enteric tablet-coating material*. Coatings of this substance disintegrate due to the hydrolytic effect of the intestinal esterases, even when the intestinal contents are acid. *In vitro* studies indicate that cellulose acetate phthalate will withstand the action of artificial gastric juices for long periods of time, but will disintegrate readily in artificial intestinal juices.

Cherry Juice

The liquid expressed from the fresh ripe fruit of *Prunus cerasus* Linné (Fam *Rosaceae*); contains not less than 1% of malic acid [$C_4H_6O_5$ = 134.09].

Preparation—Coarsely crush washed, stemmed, unpitted, sour cherries in a grinder so as to break the pits but not mash the kernels. Dissolve 0.1% of benzoic acid in the mixture, and allow it to stand at room temperature (possibly for several days) until a small portion of the filtered juice remains clear when mixed with one-half of its volume of alcohol and the resulting solution does not become cloudy within 30 min. Press the juice from the mixture and filter it.

Description—Clear liquid; aromatic, characteristicless odor; sour taste; affected by light; the color of the freshly prepared juice is red to reddish orange; pH 3 to 4; specific gravity 1.045 to 1.075.

Uses—To prepare *Cherry Syrup* (page 1301).

Carbon Tetrachloride

Methane, tetrachloro-, Tetrachloromethane

Carbon tetrachloride [56-23-5] CCl_4 (153.82).

Preparation—One method consists of catalytic chlorination of carbon disulfide.

Description—Clear, colorless liquid; characteristic odor resembling that of chloroform; specific gravity 1.588 to 1.590; boils about 77°.

Solubility—Soluble in about 2000 volumes water; miscible with alcohol, acetone, ether, chloroform or benzene.

Uses—Officially recognized as a *pharmaceutical necessity* (solvent). Formerly it was used as a cheap *anthelmintic* for the treatment of *hookworm* infections but it causes severe injury to the liver if absorbed.

Chloroform

Methane, trichloro-,

Trichloromethane [67-66-3] $CHCl_3$ (119.38); contains 99 to 99.5% of $CHCl_3$, the remainder consisting of alcohol.

Caution—Care should be taken not to vaporize it in the presence of a flame, because of the production of harmful gases (hydrogen chloride and phosgene).

Preparation—Made by the reduction of carbon tetrachloride with water and iron and by the controlled chlorination of methane.

The pure compound readily decomposes on keeping, particularly if exposed to moisture and sunlight, resulting in formation of phosgene (carbonyl chloride [$COCl_2$]) and other products. The presence of a small amount of alcohol greatly retards or prevents this decomposition; hence, the requirement that it contain 0.5 to 1% of alcohol. The alcohol combines with any phosgene forming ethyl carbonate, which is nontoxic.

Description—Clear, colorless, mobile liquid; characteristic, ethereal odor; burning, sweet taste; not flammable but its heated vapors burn with a green flame; affected by light and moisture; specific gravity 1.474 to 1.478, indicating 99 to 99.5% of $CHCl_3$; boils about 61°; not affected by acids, but is decomposed by alkali hydroxide into alkali chloride and sodium formate.

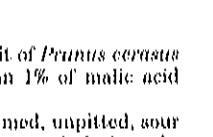
Solubility—Soluble in 210 volumes of water; miscible with alcohol, ether, benzene, solvent hexane, acetone or fixed and volatile oils.

Uses—An obsolete *inhalation anesthetic*. Although it possesses advantages of nonflammability and great potency, it rarely is used due to the serious toxic effects it produces on the heart and liver. Internally, it has been used, in small doses, as a *carminative*. Externally, it is an *irritant* and when used in liniments it may produce blisters.

It is categorized as a pharmaceutic aid. It is used as a *preservative* during the aqueous percolation of vegetable drugs to prevent bacterial decomposition in the process of manufacture. In most instances it is evaporated before the product is finished. It is an excellent solvent for alkaloids and many other organic chemicals and is used in the manufacture of these products and in chemical analyses.

Citric Acid

1,2,3-Propanetricarboxylic acid, 2-hydroxy-,



Citric acid [77-92-9] $C_6H_8O_7$ (192.12); *monohydrate* [5949-29-1] (210.14).

Preparation—Found in many plants. It formerly was obtained solely from the juice of limes and lemons and from pineapple wastes. Since about 1925 the acid has been produced largely by fermentation of sucrose solution, including molasses, by fungi belonging to the *Aspergillus niger* group, theoretically according to the following reaction

$C_{12}H_{22}O_{11}$	+ $3O_2$ →	$2H_3C_6H_5O_7$	+ $3H_2O$
Sucrose	**Oxygen**	**Citric Acid**	**Water**

but in practice there are deviations from this stoichiometric relationship.

Description—Colorless, translucent crystals, or a white, granular to fine crystalline powder; odorless; strongly acid taste; the hydrous form effloresces in moderately dry air, but is slightly deliquescent in moist air; loses its water of crystallization at about 50°; dilute aqueous solutions are subject to molding (fermentation), oxalic acid being one of the fermentation products.

Solubility—1 g in 0.5 mL of water, 2 mL of alcohol or about 30 mL of ether; freely soluble in methanol.

Uses—In the preparation of *Anticoagulant Citrate Dextrose Solution*, *Anticoagulant Citrate Phosphate Dextrose Solution*, *Citric Acid Syrup* and *effervescent salts*. It also has been used to dissolve urinary bladder calculi, and as a mild astringent.

Cocoa Butter

Cacao Butter; Theobroma Oil; Oil of Theobroma

The fat obtained from the roasted seed of *Theobroma cacao* Linné (Fam *Sterculiaceae*).

Preparation—By grinding the kernels of the "chocolate bean" and expressing the oil in powerful, horizontal hydraulic presses. The yield is about 40%. It also has been prepared by dissolving the oil from the unroasted beans by the use of a volatile solvent.

Constituents—Chemically, it is a mixture of stearin, palmitin, olein, laurin, linolein and traces of other glycerides.

Description—Yellowish, white solid; faint, agreeable odor; bland (if obtained by extraction) or chocolate-like (if obtained by pressing) taste; usually brittle below 25°; specific gravity 0.858 to 0.864 at 100°/25°; refractive index 1.454 to 1.458 at 40°.

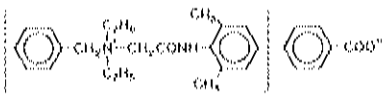
Solubility—Slightly soluble in alcohol; soluble in boiling dehydrated alcohol; freely soluble in ether or chloroform.

Uses—Valuable in pharmacy for making suppositories by virtue of its low fusing point and its property of becoming solid at a temperature just below the melting point. See *Suppositories* (page 1609). In addition to this use, it is an excellent emollient application to the skin when inflamed; it also is used in various skin creams, especially the so-called "*skin foods*." It also is used in massage.

Titanium Dioxide—page 772.

Denatonium Benzoate

Benzenemethanaminium *N*-[2-[(2,6-dimethylphenyl)amino]-2-oxoethyl]-*N,N*-diethyl-, benzoate;



Benzyldiethyl[(2,6-xylylcarbamoyl)methyl]ammonium benzoate [3734-33-6] $C_{28}H_{34}N_2O_3$ (446.59).

Preparation—2-(Diethylamino)-2',6'-xylidide is quaternized by reaction with benzyl chloride. The quaternary chloride is then treated with methanolic potassium hydroxide to form the quaternary base which, after filtering off the KCl, is reacted with benzoic acid. The starting xylidide may be prepared by condensing 2,6-xylidine with chloroacetyl chloride and condensing the resulting chloroacetoxylidide with diethylamine. US Pat 3,080,327.

Description—White, odorless, crystalline powder; an intensely bitter taste; melts about 168°.

Solubility—1 g in 20 mL of water, 2.4 mL of alcohol, 2.9 mL of chloroform or 5000 mL of ether.

Uses—A *denaturant* for ethyl alcohol.

Dextrin

British Gum; Starch Gum; Leiocum

Dextrin [9004-53-9] $(C_6H_{10}O_5)_n$.

Preparation—By the incomplete hydrolysis of starch with dilute acid, or by heating dry starch.

Description—White or yellow, amorphous powder (*white*: practically odorless; *yellow*: characteristic odor); dextrorotatory; $[\alpha]_D^{20}$ generally above 200°; does not reduce Fehling's solution; gives a reddish color with iodine.

Solubility—Soluble in 3 parts of boiling water, forming a gummy solution; less soluble in cold water.

Uses—An *excipient* and *emulsifier*.

Dextrose

Anhydrous Dextrose; Dextrose Monohydrate; Glucose; D(+)-Glucose; α-D(+)-Glucopyranose; Medicinal Glucose; Purified Glucose; Grape Sugar; Bread Sugar; Cerelose; Starch Sugar; Corn Sugar

D-Glucose monohydrate [5996-10-1] $C_6H_{12}O_6.H_2O$ (198.17); *anhydrous* [50-99-7] (180.16). A sugar usually obtained by the hydrolysis of starch. For the structure, see page 382.

Preparation—See *Liquid Glucose* (page 1321).

Description—Colorless crystals or a white, crystalline or granular powder; odorless; sweet taste; specific rotation (anhydrous) +52.5° to +53°; anhydrous dextrose melts at 146°; dextrose slowly reduces alkaline cupric tartrate TS in the cold and rapidly on heating, producing a red precipitate of cuprous oxide (difference from *sucrose*).

Solubility—1 g in 1 mL of water or 100 mL of alcohol; more soluble in boiling water or boiling alcohol.

Uses—See *Dextrose Injection* (page 800). It also is used, instead of lactose, as a supplement to milk for infant feeding.

Dichlorodifluoromethane

Methane, dichlorodifluoro-,

CCl_2F_2

Dichlorodifluoromethane [75-71-8] CCl_2F_2 (120.91).

Preparation—Carbon tetrachloride is reacted with antimony trifluoride in the presence of antimony pentafluoride.

Description—Clear, colorless gas; faint, ethereal odor; vapor pressure at 25° about 4883 torr.

Uses—A *propellant* (No 12, see page 1696).

Dichlorotetrafluoroethane

Ethane, 1,2-dichloro-1,1,2,2-tetrafluoro-,

$CClF_2CClF_2$

1,2-Dichlorotetrafluoroethane [76-14-2] $C_2Cl_2F_4$ (170.92).

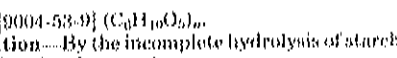
Preparation—By reacting 1,1,2-trichloro-1,2,2-trifluoroethane with antimony trifluorodichloride [SbF_3Cl_2], whereupon one of the 1-chlorine atoms is replaced by fluorine. The starting trichlorofluoroethane may be prepared from hexachloroethane by treatment with SbF_3Cl_2 (Henne AL: *Org Reactions II*: 65, 1944).

Description—Clear, colorless gas; faint, ethereal odor; vapor pressure at 25° about 1620 torr; usually contains 6 to 10% of its isomer, CF_3CCl_2 CF_3.

Uses—A *propellant* (No 114 and 114a, see page 1696).

Edetic Acid

Glycine, *N,N'*-1,2-ethanediylbis[*N*-(carboxymethyl)-,



(Ethylenedinitrilo)tetraacetic acid [60-00-4] $C_{10}H_{16}N_2O_8$ (292.24).

Preparation—Ethylenediamine is condensed with sodium monochloroacetate with the aid of sodium carbonate. An aqueous solution of the reactants is heated to about 90° for 10 hr, then cooled and acidified with HCl whereupon the acid precipitates. US Pat 2,130,505.

Description—White, crystalline powder; melts with decomposition above 220°.

Solubility—Very slightly soluble in water; soluble in solutions of alkali hydroxides.

Uses—A *pharmaceutic aid* (metal complexing agent). The acid, rather than any salt, is the form most potent in removing calcium from solution. It may be added to shed blood to prevent clotting. It also is used in pharmaceutical analysis and the removal or inactivation of unwanted ions in solution. Salts of the acid are known as edetates. See *Edetate Calcium Disodium* (page 824) and *Edetate Disodium* (page 825).

Ethylcellulose

Cellulose ethyl ether [9004-57-3]; an ethyl ether of cellulose containing 44 to 51% of ethoxy groups. The *medium-type* viscosity grade contains less than 46.5% ethoxy groups; the *standard-type* viscosity grade contains 46.5% or more ethoxy groups.

Preparation—By the same general procedure described on page 1306 for *Methylcellulose* except that ethyl chloride or ethyl sulfate is employed as the alkylating agent. The 45 to 50% of ethoxy groups in the official ethylcellulose corresponds to from 2.25 to 2.61 ethoxy groups/$C_6H_{10}O_5$ unit, thus representing from 75 to 87% of the maximum theoretical ethoxylation, which is 3 ethoxy groups/$C_6H_{10}O_5$ unit.

Description—Free-flowing, white to light tan powder; forms films that have a refractive index of about 1.47; aqueous suspensions are neutral to litmus.

Solubility—The medium-type is freely soluble in tetrahydrofuran, methyl acetate, chloroform or mixtures of aromatic hydrocarbons with alcohol; the standard-type is freely soluble in alcohol, methanol, toluene, chloroform or ethyl acetate; both types are insoluble in water, glycerin or propylene glycol.

Uses—A *pharmaceutic aid* as a tablet binder and for film-coating tablets and drug particles.

Gelatin—page 1306.

Liquid Glucose

Glucose; Starch Syrup; Corn Syrup

A product obtained by the incomplete hydrolysis of starch. It consists chiefly of dextrose [D(+)-glucose, $C_6H_{12}O_6$ = 180.16] dextrins, maltose and water.

Preparation—Commercially by the action of very weak H_2SO_4 or HCl on starch.

One of the processes for its manufacture is as follows: The starch, usually from corn, is mixed with 5 times its weight of water containing less than 1% of HCl, the mixture is heated to about 45° and then transferred to a suitable reaction vessel into which steam is passed

under pressure until the temperature reaches 120°. The temperature is maintained at this point for about 1 hr, or until tests show complete disappearance of starch. The mass is then heated to volatilize most of the hydrochloric acid, sodium carbonate or calcium carbonate is added to neutralize the remaining traces of acid, the liquid is filtered, then decolorized in charcoal or bone-black filters, as is done in sugar refining and finally concentrated in vacuum to the desired consistency.

When made by the above process, it contains about 30 to 40% of dextrose mixed with about an equal proportion of dextrin, together with small amounts of other carbohydrates, notably maltose. By varying the conditions of hydrolysis, the relative proportions of the sugars also vary.

If the crystallizable dextrose is desired, the conversion temperature is higher and the time of conversion longer. The term "glucose," as customarily used in the chemical or pharmaceutical literature, usually refers to dextrose, the crystallizable product.

The name "grape sugar" sometimes is applied to the solid commercial form of dextrose because the principal sugar of the grape is dextrose, although the fruit has never been used as a source of the commercial supply.

Description—Colorless or yellowish, thick, syrupy liquid; odorless, or nearly so; sweet taste; differs from sucrose in that it readily reduces hot alkaline cupric tartrate TS, producing a red precipitate of cuprous oxide.
Solubility—Miscible with water; sparingly soluble in alcohol.

Uses—As an ingredient of *Cocoa Syrup* (page 1301), as a tablet binder and coating agent, and as a diluent in pilular extracts; it has replaced glycerin in many pharmaceutical preparations. It is sometimes given *per rectum* as a *food* in cases where feeding by stomach is impossible. It should not be used in the place of dextrose for intravenous injection.

Hydrochloric Acid

Chlorhydric Acid; Muriatic Acid; Spirit of Salt

Hydrochloric acid [7647-01-0] HCl (36.46); contains 36.5 to 38.0%, by weight, of HCl.

Preparation—By the interaction of NaCl and H_2SO_4 or by combining chlorine with hydrogen. It is obtained as a by-product in the manufacture of sodium carbonate from NaCl by the Leblanc process in which common salt is decomposed with H_2SO_4. HCl is also a by-product in the electrolytic production of NaOH from NaCl.

Description—Colorless, fuming liquid; pungent odor; fumes and odor disappear when it is diluted with 2 volumes of water; strongly acid to litmus even when highly diluted; specific gravity about 1.18.
Solubility—Miscible with water or alcohol.

Uses—Officially classified as a pharmaceutic aid that is used as an acidifying agent. It is used in preparing *Diluted Hydrochloric Acid* (page 793).

Hypophosphorous Acid

Phosphinic acid

Hypophosphorous acid [6303-21-5] HPH_2O_2 (66.00); contains 30 to 32% by weight, of H_3PO_2.

Preparation—By reacting barium or calcium hypophosphite with sulfuric acid or by treating sodium hypophosphite with an ion-exchange resin.

Description—Colorless or slightly yellow, odorless liquid; solution is acid to litmus even when highly diluted; specific gravity about 1.13.
Solubility—Miscible with water or alcohol.
Incompatibilities—Oxidized on exposure to air and by nearly all *oxidizing agents*. *Mercury, silver* and *bismuth salts* are reduced partially to the metallic state as evidenced by a darkening in color. *Ferric compounds* are changed to ferrous.

Uses—An *antioxidant* in pharmaceutical preparations.

Isopropyl Myristate

Tetradecanoic acid, 1-methylethyl ester

$CH_3(CH_2)_{12}COOCH(CH_3)_2$

Isopropyl myristate [110-27-0] $C_{17}H_{34}O_2$ (270.45).

Preparation—By reacting myristoyl chloride with 2-propanol with the aid of a suitable dehydrochlorinating agent.

Description—Liquid of low viscosity; practically colorless and odorless; congeals about 5° and decomposes at 208°; withstands oxidation and does not become rancid readily.
Solubility—Soluble in alcohol, acetone, chloroform, ethyl acetate, toluene, mineral oil, castor oil or cottonseed oil; practically insoluble in water, glycerin or propylene glycol; dissolves many waxes, cholesterol or lanolin.

Uses—*Pharmaceutic aid* used in cosmetics and topical medicinal preparations as an emollient, lubricant and to enhance absorption through the skin.

Kaolin—see page 790.

Lactic Acid

Propanoic acid, 2-hydroxy-, 2-Hydroxypropionic Acid; Propanoloic Acid; Milk Acid

$CH_3CH(OH)COOH$

Lactic acid [50-21-5] $C_3H_6O_3$ (90.08); a mixture of lactic acid and lactic acid lactate ($C_6H_{10}O_5$) equivalent to a total of 85 to 90%, by weight, of $C_3H_6O_3$.

Discovered by Scheele in 1780, it is the acid formed in the souring of milk, hence the name *lactic*, from the Latin name for milk. It results from the decomposition of the lactose (milk sugar) in milk.

Preparation—A solution of glucose or of starch previously hydrolyzed with diluted sulfuric acid is inoculated, after the addition of suitable nitrogen compounds and mineral salts, with *Bacillus lactis*. Calcium carbonate is added to neutralize the lactic acid as soon as it is formed, otherwise the fermentation stops when the amount of acid exceeds 0.5%. When fermentation is complete, as indicated by failure of the liquid to give a test for glucose, the solution is filtered, concentrated and allowed to stand. The calcium lactate that crystallizes is decomposed with dilute sulfuric acid and filtered with charcoal. The lactic acid in the filtrate is extracted with ethyl or isopropyl ether, the ether is distilled off and the aqueous solution of the acid concentrated under reduced pressure.

Description—Colorless or yellowish, nearly odorless, syrupy liquid; acid to litmus; absorbs water on exposure to moist air; when a dilute solution is concentrated to above 50%, lactic acid lactate begins to form; in the official acid the latter amounts to about 12 to 15%; specific gravity about 1.20; decomposes when distilled under normal pressure but may be distilled without decomposition under reduced pressure.
Solubility—Miscible with water, alcohol or ether; insoluble in chloroform.

Uses—In the preparation of *Sodium Lactate Injection* (page 821). It also is used in babies' milk formulas, as an acidulant in food preparations, and in 1 to 2% concentration in some spermatocidal jellies. A 10% solution is used as a bactericidal agent on the skin of neonates. It is corrosive to tissues on prolonged contact. A 16.7% solution in flexible collodion is used to remove warts and small cutaneous tumors.

Lactose

D-Glucose, 4-*O*-β-D-galactopyranosyl-, Milk Sugar

Lactose [63-42-3] $C_{12}H_{22}O_{11}$ (342.30); *monohydrate* [10039-26-6] (360.31); a sugar obtained from milk.

For the structural formula, see page 382.

Preparation—From skim milk, to which is added diluted HCl to precipitate the casein. After removal of the casein by filtration, the reaction of the whey is adjusted to a pH of about 6.2 by addition of lime and the remaining albuminous matter is coagulated by heating; this is filtered out and the liquid set aside to crystallize. Animal charcoal is used to decolorize the solution in a manner similar to that used in purifying sucrose.

Another form of lactose, known as β-lactose, also is available on the market. It differs in that the D-glucose moiety is β instead of α. It is reported that this variety is sweeter and more soluble than ordinary lactose and for that reason is preferable in pharmaceutical manufacturing where lactose is used. Chemically, β-lactose does not appear to differ from ordinary α-lactose. It is manufactured in the same way as α-lactose up to the point of crystallization, then the

solution is heated to a temperature above 93.5°, this being the temperature at which the α form is converted to the β variety. The β form occurs only as an anhydrous sugar whereas the α variety may be obtained either in the anhydrous form or as a monohydrate.

Description—White or creamy white, hard, crystalline masses or powder; odorless; faintly sweet taste; stable in air, but readily absorbs odors; pH (1 in 10 solution) 4.0 to 6.5; specific rotation +54.8 to +55.5°.
Solubility—1 g in 5 mL of water or 2.6 mL of boiling water; very slightly soluble in alcohol; insoluble in chloroform or ether.

Uses—A *diluent* largely used in medicine and pharmacy. It is generally an ingredient of the medium used in penicillin production. It is used extensively as an addition to milk for infant feeding.

Magnesium Chloride

Magnesium chloride hexahydrate [7791-18-6] $MgCl_2.6H_2O$ (203.30); *anhydrous* [7786-30-3] (95.21).
Preparation—By treating magnesite or other suitable magnesium minerals with HCl.

Description—Colorless, odorless, deliquescent flakes or crystals, which lose water when heated to 100° and loose HCl when heated to 110°; pH (1 in 20 solution in carbon dioxide free water) 4.5 to 7.
Solubility—Very soluble in water; freely soluble in alcohol.

Uses—*Electrolyte replenisher; pharmaceutical necessity* for hemodialysis and peritoneal dialysis fluids.

Magnesium Stearate

Octadecanoic acid, magnesium salt

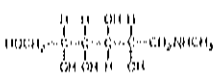
Magnesium stearate [557-04-0]. A compound of magnesium with a mixture of solid organic acids obtained from fats, and consists chiefly of variable proportions of magnesium stearate and magnesium palmitate. It contains the equivalent of 6.8 to 8.0% of MgO (40.30).

Description—Fine, white, bulky powder; faint, characteristic odor; unctuous, adheres readily to the skin and free from grittiness.
Solubility—Insoluble in water, alcohol or ether.

Uses—A *pharmaceutical necessity* (*lubricant*) in the manufacture of compressed tablets.

Meglumine

D-Glucitol, 1-deoxy-1-(methylamino)-,



1-Deoxy-1-(methylamino) D-glucitol [6284-40-8] $C_7H_{17}NO_5$ (195.21).
Preparation—By treating glucose with hydrogen and methylamine under pressure and in the presence of Raney nickel.

Description—White to faintly yellowish white, odorless crystals or powder; melts about 130°.
Solubility—Freely soluble in water; sparingly soluble in alcohol.

Uses—In forming salts of certain pharmaceuticals, surface-active agents and dyes. See *Diatrizoate Meglumine Injections* (page 1276), *Iodipamide Meglumine Injection* (page 1276) and *Iothalamate Meglumine Injection* (page 1277).

Light Mineral Oil

Light Liquid Petrolatum NF XIV; Light Liquid Paraffin; Light White Mineral Oil

A mixture of liquid hydrocarbons obtained from petroleum. It may contain a suitable stabilizer.

Description—Colorless, transparent, oily liquid, free, or nearly free, from fluorescence; odorless and tasteless when cold, and develops not more than a faint odor of petroleum when heated; specific gravity 0.818 to 0.880; kinematic viscosity not more than 35.5 centistokes at 40°.
Solubility—Insoluble in water or alcohol; miscible with most fixed oils, but not with castor oil; soluble in volatile oils.

Uses—Officially recognized as a *vehicle*. Once it was used widely as a vehicle for nose and throat medications; such uses are now considered dangerous because of the possibility of lipoid pneumonia. It sometimes is used to cleanse dry and inflamed skin areas and to facilitate removal of dermatological preparations from the skin. It should never be used for internal administration because of "leakage." See *Mineral Oil* (page 789).

Nitric Acid

Nitric acid [7697-37-2] HNO_3 (63.01); contains about 70%, by weight, of HNO_3.
Preparation—May be prepared by treatment of sodium nitrate (Chile saltpeter) with sulfuric acid, but usually produced by catalytic oxidation of ammonia.

Description—Highly corrosive fuming liquid; characteristic, highly irritating odor; stains animal tissues yellow; boils about 120°; specific gravity about 1.41.
Solubility—Miscible with water.

Uses—*Pharmaceutic aid* (acidifying agent).

Nitrogen

Nitrogen [7727-37-9] N_2 (28.01); contains not less than 99%, by volume, of N_2.
Preparation—By the fractional distillation of liquified air.
Uses—A diluent for medicinal gases. Pharmaceutically, is employed to replace air in the containers of substances which would be affected adversely by air oxidation. Examples include its use with fixed oils, certain vitamin preparations and a variety of injectable products. It also is used as a propellant.

Persic Oil

Apricot Kernel Oil; Peach Kernel Oil

The oil expressed from the kernels of varieties of *Prunus armeniaca* Linné (Apricot Kernel Oil), or from the kernels of varieties of *Prunus persica* Sieb et Zucc (Peach Kernel Oil) (Fam *Rosaceae*).

Description—Clear, pale straw-colored or colorless, almost odorless, oily liquid with a bland taste; specific gravity 0.910 to 0.923; not turbid at temperatures above 15°.
Solubility—Slightly soluble in alcohol; miscible with ether, chloroform, benzene or solvent hexane.

Uses—A *vehicle*. It also is used in preparing cold creams.

Phenol

Carbolic Acid

C_6H_5OH

Phenol [108-95-2] C_6H_6O (94.11).
Preparation—For many years made only by distilling crude carbolic acid from coal tar and separating and purifying the distillate by repeated crystallizations, it now is prepared synthetically.
A more recent process uses chlorobenzene as the starting point in the manufacture. The chlorobenzene is produced in a vapor phase reaction, with benzene, HCl and oxygen over a copper catalyst, followed by hydrolysis with steam to yield HCl and phenol (which is recovered).

Description—Colorless to light pink, interlaced, or separate, needle-shaped crystals, or a white or light pink, crystalline mass; characteristic odor; when undiluted, it whitens and cauterizes the skin and mucous membranes; when gently heated, phenol melts, forming a highly refractive liquid; liquefied by the addition of 10% of water; vapor is flammable; gradually darkens on exposure to light and air; specific gravity 1.07; boils at 182°; congeals not lower than 39°.
Solubility—1 g in 15 mL of water; very soluble in alcohol, glycerin, chloroform, ether or fixed and volatile oils; sparingly soluble in mineral oil.
Incompatibilities—Produces a liquid or soft mass when triturated with *camphor, menthol, acetanilid, acetophenetidin, aminopyrine, antipyrine, ethyl aminobenzoate, methenamine, phenyl salicylate, resorcinol, terpin hydrate, thymol* and several other substances including some *alkaloids*. It also softens *cocoa butter* in suppository mixtures.
It is soluble in about 15 parts of water; stronger solutions may be

obtained by using as much glycerin as phenol. Only the crystallized form is soluble in fixed oils and liquid petroleum, the liquefied form is not all soluble due to its content of water. *Albumin* and *gelatin* are precipitated by it. *Collodion* is coagulated by the precipitation of pyroxylin. Traces of *iron* in various chemicals such as *alum*, *borax*, etc, may produce a green color.

Uses—A *caustic*, *disinfectant*, *topical anesthetic* and pharmaceutical necessity as a *preservative* for injections, etc. At one time widely used as a germicide and still the standard against which other antiseptics are compared, it has few legitimate uses in modern medicine. Nevertheless, it is still used in several proprietary antiseptic mouthwashes, hemorrhoidal preparations and burn remedies. In full strength, a few drops of the liquefied form may be used to cauterize small wounds, dogs bites, snake bites, etc. It commonly is employed as an *antipruritic*, either in the form of phenolated calamine lotion (1%), phenol ointment (2%) or a simple aqueous solution (0.5 to 1%). It has been used for sclerosing hemorrhoids, but more effective and safer drugs are available. A 5% solution in glycerin is used in simple earache. Crude carbolic acid is an effective, economical agent for disinfecting excrement. It is of some therapeutic value as a *fungicide*, but more effective and less toxic agents are available. If accidentally spilled, it should be removed promptly from the skin by swabbing with alcohol.

Liquefied Phenol [Liquefied Carbolic Acid is phenol maintained in a liquid condition by the presence of 10.0% of water. It contains not less than 89.0%, by weight, of C_6H_6O. *Note—When it is to be mixed with a fixed oil, mineral oil or white petrolatum, use the crystalline Phenol, not Liquefied Phenol. Preparation:* Melt phenol (a convenient quantity) by placing the unstoppered container in a steam bath and applying heat gradually. Transfer the liquid to a tared vessel, weigh, add 1 g of purified water for each 9 g of phenol, and mix thoroughly. *Description:* Colorless liquid, which may develop a red tint upon exposure to air and light; characteristic, somewhat aromatic odor; when undiluted it cauterizes and whitens the skin and mucous membranes; specific gravity about 1.065; when it is subjected to distillation, the boiling temperature does not rise above 182°, which is the boiling temperature of phenol; partially solidifies at about 15°. *Solubility:* Miscible with alcohol, ether or glycerin; a mixture of liquefied phenol and an equal volume of glycerin is miscible with water. *Uses:* A formulation which facilitates the dispensing of concentrated phenol. Its therapeutic uses are described above under *Phenol*. It is a *pharmaceutical necessity for Phenolated Calamine Lotion* (page 762).

Phenyl Salicylate—RPS-15, page 1269.

Phosphoric Acid

Orthophosphoric Acid; Syrupy Phosphoric Acid; Concentrated Phosphoric Acid

Phosphoric acid [7664-38-2] H_3PO_4 (98.00); contains 85 to 88%, by weight, of H_3PO_4.

Preparation—Phosphorus is converted to phosphorus pentoxide [P_2O_5] by exposing it to a current of warm air, then the P_2O_5 is treated with water to form phosphoric acid. The conversion of the phosphorus to the pentoxide takes place while the phosphorus, distilling from the phosphorus manufacturing operation, is in the vapor state.

Description—Colorless, odorless liquid of a syrupy consistency; specific gravity about 1.71.

Solubility—Miscible with water or alcohol, with the evolution of heat.

Uses—To make the diluted acid and as a weak acid in various pharmaceutical preparations. Industrially, it is used in dental cements and in beverages as an acidulant.

Diluted Phosphoric Acid [Dilute Phosphoric Acid] contains, in each 100 mL, 9.5 to 10.5 g of H_3PO_4 (98.00). *Preparation:* Mix phosphoric acid (69 mL) and purified water (qs) to make 1000 mL. *Description:* Clear, colorless, odorless liquid; specific gravity about 1.057. Miscible with water or alcohol. *Uses:* A *pharmaceutical necessity*. It also has been employed in *lead poisoning* and in other conditions in which it is desired to administer large amounts of phosphate and at the same time produce a mild acidosis. It has been given in the dose of 60 mL a day (5 mL/hour) under carefully controlled conditions.

Potassium Metaphosphate

Metaphosphoric acid (HPO_3), potassium salt

Potassium metaphosphate [7790-53-6] KPO_3 (118.07); a straight-chain polyphosphate, having a high degree of polymerization; contains the equivalent of 59 to 61% of P_2O_5.

Preparation—By thermal dehydration of monopotassium phosphate (KH_2PO_4).

Description—White, odorless powder.

Solubility—Insoluble in water; soluble in dilute solutions of sodium salts.

Uses—*Pharmaceutic aid* (buffering agent).

Monobasic Potassium Phosphate

Phosphoric acid, monopotassium salt; Potassium Biphosphate; Potassium Acid Phosphate; Potassium Dihydrogen Phosphate; Sörensen's Potassium Phosphate

Monopotassium phosphate [7778-77-0] KH_2PO_4 (136.09).

Preparation—H_3PO_4 is reacted with an equimolar quantity of KOH and the solution is evaporated to crystallization.

Description—Colorless crystals or a white, granular or crystalline powder; odorless and stable in air; pH (1 in 100 solution) about 4.5.

Solubility—Freely soluble in water; practically insoluble in alcohol.

Uses—A component of various buffer solutions. Medicinally, it has been used as a urinary acidifier.

Pumice

Pumex

A substance of volcanic origin, consisting chiefly of complex silicates of aluminum, potassium and sodium.

Description—Very light, hard rough, porous, grayish masses or a gritty, grayish powder of several grades of fineness; odorless, tasteless and stable in the air.

Three powders are available:

Pumice Flour or *Superfine Pumice*—Not less than 97% passes through a No 200 standard mesh sieve.

Fine Pumice—Not less than 95% passes through a No 150 standard mesh sieve, and not more than 75% passes through a No 200 standard mesh sieve.

Coarse Pumice—Not less than 95% passes through a No 60 standard mesh sieve, and not more than 5% passes through a No 200 standard mesh sieve.

Solubility—Insoluble in water and is not attacked by acids or alkali hydroxide solutions.

Uses—A *filtering* and *distributing medium* for pharmaceutical preparations. Because of its grittiness the powdered form is used in certain types of soaps and cleaning powders and also as a *dental abrasive*.

Pyroxylin

Cellulose, nitrate; Soluble Guncotton

Pyroxylin [9004-70-0]; a product obtained by the action of a mixture of nitric and sulfuric acids on cotton, and consists chiefly of cellulose tetranitrate [$(C_{12}H_{16}N_4O_{18})_n$].

Note—The commercially available form is moistened with about 30% of alcohol or other suitable solvent. The alcohol or solvent must be allowed to evaporate to yield the dried substance described in the Pharmacopeia.

Preparation—Schönbein, in 1846, found that nitric acid acts on cotton and produces a soluble compound. It subsequently was proved that this substance belongs to a series of closely related nitrates in which the nitric acid radical replaces the hydroxyl of the cellulose formula. This usually is indicated by taking the double empirical formula for cellulose $C_{12}H_{20}O_{10}$ and indicating replacement of four of the OH groups thus

$$\underset{\text{Cellulose}}{C_{12}H_{20}O_{10}} + 4HNO_3 \rightarrow \underset{\substack{\text{Cellulose}\\\text{Tetranitrate}}}{C_{12}H_{16}O_6(NO_3)_4} + 4H_2O$$

The compound used in preparing collodion is a varying mixture of the di-, tri-, tetra- and pentanitrates, but is mainly tetranitrate. The hexanitrate is the true explosive guncotton, and is insoluble in ether, alcohol, acetone or water.

Description—Light yellow, matted mass of filaments, resembling raw cotton in appearance, but harsh to the touch; *exceedingly flammable* burning, when unconfined, very rapidly and with a luminous flame; when kept in well-closed bottles and exposed to light, it is decomposed with the evolution of nitrous vapors, leaving a carbonaceous residue.

Solubility—Insoluble in water; dissolves slowly but completely in 25 parts of a mixture of 3 volumes of ether and 1 volume of alcohol; soluble in acetone or glacial acetic acid and precipitated from these solutions by water.

Uses—A *pharmaceutical necessity* for *Collodion* (RPS-16, page 717).

Rosin

Resina; Colophony; Georgia Pine Rosin; Yellow Pine Rosin

A solid resin obtained from *Pinus palustris* Miller, and from other species of *Pinus* Linné (Fam *Pinaceae*).

Constituents—American rosin contains *sylvic acid* [$C_{20}H_{30}O_2$], α-, β- and γ-*abietic acids* [$C_{20}H_{30}O_2$], γ-*pinic acid* (from which α- and β-pinic acids are gradually formed) and *resene*. Some authorities also include *pimaric acid* [$C_{20}H_{30}O_2$] as a constituent. French rosin is called *galipot*.

Description—Sharply angular, translucent, amber-colored fragments, frequently covered with a yellow dust; fracture brittle at ordinary temperatures, shiny and shallow-conchoidal; odor and taste are slightly terebinthinate; easily fusible and burns with a dense, yellowish smoke, specific gravity 1.07 to 1.09.

Solubility—Insoluble in water; soluble in alcohol, ether, benzene, glacial acetic acid, chloroform, carbon disulfide, dilute solutions of sodium hydroxide and potassium hydroxide or some volatile and fixed oils.

Uses—A pharmaceutical necessity for *Zinc-Eugenol Cement* (page 1328). Formerly, and to some extent still, used as a component of plasters, cerates and ointments, to which it adds adhesive qualities.

Purified Siliceous Earth

Purified Kieselguhr; Purified Infusorial Earth; Diatomaceous Earth; Diatomite

A form of silica [SiO_2] [7631-86-9] consisting of the frustules and fragments of diatoms, purified by boiling with acid, washing and calcining.

Occurrence and Preparation—Large deposits of this substance are found in Virginia, Maryland, Nevada, Oregon and California, usually in the form of masses of rocks, hundreds of feet in thickness. Under the microscope it is seen to consist largely of the minute siliceous frustules of diatoms. It must be purified carefully in a manner similar to that directed for *Talc* (page 1327), and thoroughly calcined. The latter treatment destroys the bacteria which are present in large quantities in the native earth.

Description—Very fine, white, light-gray or pale-buff mixture of amorphous powder and lesser amounts of crystalline polymorphs, including quartz and cristobalite; gritty, readily absorbs moisture and retains about four times its weight of water without becoming fluid.

Solubility—Insoluble in water, acids or dilute solutions of alkali hydroxides.

Uses—Introduced into the USP as a distributing and *filtering medium* for aromatic waters; also suitable for filtration of elixirs. Like talc, it does not absorb active constituents.

Colloidal Silicon Dioxide

Silica [7631-86-9] SiO_2 (60.08); a submicroscopic fumed silica prepared by the vapor-phase hydrolysis of a silicon compound.

Description—Light, white, nongritty powder of extremely fine particle size (about 15 nm).

Solubility—Insoluble in water or acids (except hydrofluoric); dissolved by hot solutions of alkali hydroxides.

Uses—A *tablet diluent* and as a *suspending* and *thickening agent* in pharmaceutical preparations.

Soda Lime

A mixture of calcium hydroxide and sodium or potassium hydroxide or both.

It may contain an indicator that is inert toward anesthetic gases such as ether, cyclopropane and nitrous oxide, and that changes color when the soda lime no longer can absorb carbon dioxide.

Description—White or grayish white granules; if an indicator is added, it may have a color; absorbs carbon dioxide and water on exposure to air.

Uses—Neither a therapeutic nor a pharmaceutical agent. It is a *reagent for the absorption of carbon dioxide* in anesthesia machines, oxygen therapy and metabolic tests. Because of the importance of the proper quality for these purposes it has been made official and standardized.

Sodium Borate

Sodium Tetraborate; Sodium Pyroborate; Sodium Biborate

Borax [1303-96-4] $Na_2B_4O_7.10H_2O$ (381.37); anhydrous [1330-43-4] $Na_2B_4O_7$ (201.22).

Preparation—Found in immense quantities in California as a crystalline deposit. The earth, which is strongly impregnated with borax, is lixiviated; the solution is evaporated and crystallized.

Calcium borate, or *cotton bulls*, also occurs in the borax deposits of California, and sodium borate is obtained from it by double decomposition with sodium carbonate.

Description—Colorless, transparent crystals, or a white, crystalline powder; odorless; the crystals often are coated with white powder due to efflorescence; solution alkaline to litmus and phenolphthalein; pH about 9.5.

Solubility—1 g in 16 mL of water, 1 mL of glycerin or 1 mL of boiling water; insoluble in alcohol.

Incompatibilities—Precipitates many *metals* as insoluble borates. In aqueous solution it is alkaline and precipitates *aluminum salts* as aluminum hydroxide, *iron salts* as a basic borate and ferric hydroxide and *zinc sulfate* as zinc borate and a basic salt. *Alkaloids* are precipitated from solutions of their salts. Approximately equal weights of *glycerin* and boric acid react to produce a decidedly acid derivative generally called glyceroboric acid. Thus, the addition of glycerin to a mixture containing it overcomes incompatibilities arising from an alkaline reaction.

Uses—As a pharmaceutical necessity, it is used as an alkalizing agent and as a buffer for alkaline solutions. Its alkalizing properties provide the basis for its use in denture adhesives and its buffering action for its use in eyewash formulations.

Sodium Carbonate

Carbonic acid, disodium salt, monohydrate;
Monohydrated Sodium Carbonate USP XVII

Disodium carbonate monohydrate [5968-11-6] $Na_2CO_3.H_2O$ (124.00); *anhydrous* [497-19-8] (105.99).

Preparation—The initial process for its manufacture was devised by Leblanc, a French apothecary, in 1784, and consists of two steps: first, the conversion of common salt [NaCl] into sodium sulfate by heating it with sulfuric acid and, second, the decomposition of the sulfate by calcium carbonate (limestone) and charcoal (coal) at a high temperature to yield this salt and calcium sulfide. The carbonate then is leached out with water.

It currently is prepared by the electrolysis of sodium chloride, whereby sodium and chlorine are produced, the former reacting with water to produce sodium hydroxide and this solution treated with carbon dioxide to produce the salt. The process is used most extensively in localities where electric power is very cheap.

The monohydrated form is made by crystallizing a concentrated solution of this salt at a temperature above 35° (95°F), and stirring the liquid so as to produce small crystals. It contains about 15% of water of crystallization.

Soda ash is a term designating a commercial quality of the anhydrous salt. Its annual production is very large, and it has a wide variety of applications, among which are the manufacture of glass, soap and sodium salts; it also is used for washing fabrics.

Washing soda, or *sal soda*, is the salt with 10 molecules of water. It is in the form of colorless crystals which rapidly effloresce in the air.

Description—Colorless crystals or a white, crystalline powder; stable in air under ordinary conditions; when exposed to dry air above 50° it effloresces, and at 100° becomes anhydrous; decomposed by weak acids forming the salt of the acid and liberating carbon dioxide; aqueous solution alkaline to indicators (pH about 11.5).

Solubility—1 g in 3 mL of water or 1.5 mL of boiling water; insoluble in alcohol.

Incompatibilities—*Acids, acid salts* and *acidic preparations* cause its decomposition. Most *metals* are precipitated as carbonates, hydroxides or basic salts. *Alkaloids* are precipitated from solutions of their salts.

Uses—Occasionally, for dermatitides topically as a lotion; it has been used as a mouthwash and a vaginal douche. It is used in the preparation of the sodium salts of many acids. The USP recognizes it as a pharmaceutic aid used as an alkalizing agent.

Sodium Hydroxide

Caustic Soda, Soda Lye

Sodium hydroxide [1310-73-2] NaOH (40.00); includes not more than 3% of Na_2CO_3 (105.99).

Caution—Exercise great care in handling it, as it rapidly destroys tissues.

Preparation—By treating sodium carbonate with milk of lime, or by the electrolysis of a solution of sodium chloride as explained under *Potassium Hydroxide* (page 767). It now is produced largely by the latter process. See also *Sodium Carbonate*, above.

Description—White, or nearly white, fused masses, small pellets, flakes, sticks and other forms; hard and brittle and shows a crystalline fracture; exposed to the air, it rapidly absorbs carbon dioxide and moisture; melts at about 318°; specific gravity 2.13; when dissolved in water or alcohol, or when its solution is treated with an acid, much heat is generated; aqueous solutions, even when highly diluted, are strongly alkaline.

Solubility—1 g in 1 mL of water; freely soluble in alcohol or glycerin.

Incompatibilities—Exposed to air, it absorbs *carbon dioxide* and is converted to sodium carbonate. With *fats* and *fatty acids* it forms soluble soaps; with *resins* it forms insoluble soaps. See *Potassium Hydroxide* (page 767).

Uses—Too alkaline to be of medicinal value but occasionally used in veterinary practice as a caustic. It is used extensively in pharmaceutical processes as an alkalizing agent and is generally preferred to potassium hydroxide because it is less deliquescent, and less expensive; in addition, less of it is required since 40 parts of it are equivalent to 56 parts of KOH. It is a pharmaceutical necessity in the preparation of *Glycerin Suppositories* (page 765).

Sodium Stearate

Octadecanoic acid, sodium salt

Sodium stearate [822-16-2] $C_{18}H_{35}NaO_2$ (306.47) consists chiefly of sodium stearate and sodium palmitate [$C_{16}H_{31}NaO_2$ = 278.41].

Preparation—Stearic acid is reacted with an equimolar portion of NaOH.

Description—Fine, white powder, soapy to the touch; usually has a slight, tallow-like odor; affected by light; solutions are alkaline to phenolphthalein TS.

Solubility—Slowly soluble in cold water and cold alcohol; readily soluble in hot water and hot alcohol.

Uses—Officially, a pharmaceutic aid used as an emulsifying and stiffening agent. It is an ingredient of glycerin suppositories. In dermatological practice it has been used topically in sycosis and other skin diseases.

Starch

Corn Starch; Wheat Starch; Potato Starch

Starch [9005-25-8]; consists of the granules separated from the mature grain of corn [*Zea mays* Linné (Fam *Gramineae*)] or of wheat [*Triticum aestivum* Linné] (Fam *Gramineae*), or from tubers of the potato [*Solanum tuberosum* Linné (Fam *Solanaceae*)].

Preparation—In making starch from corn, the germ is separated mechanically and the cells softened to permit escape of the starch granules. This generally is done by permitting it to become sour and decomposed, stopping the fermentation before the starch is affected. On the small scale, it may be made from wheat flour by making a stiff ball of dough and kneading it while a small stream of water trickles upon it. It is carried off with the water, while the *gluten* remains as a soft, elastic mass; the latter may be purified and used for various purposes to which gluten is applicable. Commercially, its quality largely depends on the purity of the water used in its manufacture. It may be made from potatoes by first grating them, and then washing the soft mass upon a sieve, which separates the cellular substances and permits the starch granules to be carried through. It then must be washed thoroughly by decantation, and the quality of this starch also depends largely on the purity of the water that is used in washing it.

Description—Irregular, angular, white masses or fine powder; odorless; slight, characteristic taste. *Corn starch:* Polygonal, rounded or spheroidal granules up to about 35 μm in diameter which usually have a circular or several-rayed central cleft. *Wheat starch:* Simple lenticular granules 20 to 50 μm in diameter and spherical granules 5 to 10 μm in diameter; striations faintly marked and concentric. *Potato starch:* Simple granules, irregularly ovoid or spherical, 30 to 100 μm in diameter, and subspherical granules 10 to 30 μm in diameter; striations well-marked and concentric.

Solubility—Insoluble in cold water or alcohol; when it is boiled with about 20 times its weight of hot water for a few minutes and then cooled, a translucent, whitish jelly results; aqueous suspension neutral to litmus.

Uses—Has absorbent and demulcent properties. It is used as a dusting powder and in various dermatological preparations; also as a pharmaceutic aid (filler, binder and disintegrant). *Note—Starches obtained from different botanical sources may not have identical properties with respect to their use for specific pharmaceutical purposes, eg, as a tablet-disintegrating agent. Therefore, types should not be interchanged unless performance equivalency has been ascertained.*

Under the title *Pregelatinized Starch* the NF recognizes starch that has been processed chemically or mechanically to rupture all or part of the granules in the presence of water, and subsequently dried. Some types may be modified to render them compressible and flowable.

Storax

Liquid Storax; Styrax; Sweet Gum; Prepared Storax

A balsam obtained from the trunk of *Liquidambar orientalis* Miller, known in commerce as Levant Storax, or of *Liquidambar styraciflua* Linné, known in commerce as American Storax (Fam *Hamamelidaceae*).

Constituents—The following occur in both varieties: *styracin* (*cinnamyl cinnamate*), *styrol* (*phenylethylene*, C_8H_8), *α-* and *β-storesin* (the cinnamic acid ester of an alcohol called *storesinol*), *phenylpropyl cinnamate*, *free cinnamic acid* and *vanillin*. In addition to these, Levant storax contains *ethyl cinnamate*, *benzyl cinnamate*, *free storesinol*, *isocinnamic acid*, *ethylvanillin*, *styrogenin* and *styrocamphene*. This variety yields from 0.5 to 1% of *volatile oil*; from this have been isolated *styrocamphene*, *vanillin*, the cinnamic acid esters of *ethyl*, *phenylpropyl*, *benzyl* and *cinnamyl alcohols*, *naphthalene* and *styrol*.

The American variety contains, in addition to the aforementioned substances common to both varieties, *styaresin* (the cinnamic acid ester of the alcohol *styresinol*, an isomer of storesinol) and *styresinolic acid*. It yields up to 7% of a dextrorotatory volatile oil, the composition of which has not been investigated completely; styrol and traces of vanillin have been isolated from it.

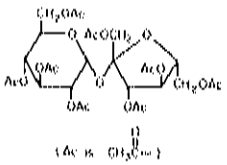
Description—Semiliquid, grayish to grayish brown, sticky, opaque mass, depositing on standing a heavy dark brown layer (Levant storax); or a semisolid, sometimes a solid mass, softened by gently warming (American storax); transparent in thin layers; characteristic odor and taste; more dense than water.

Solubility—Insoluble in water, but soluble, usually incompletely, in an equal weight of warm alcohol; soluble in acetone, carbon disulfide or ether, some insoluble residue usually remaining.

Uses—An *expectorant* but is used chiefly as a local remedy, especially in combination with benzoin; eg, it is an ingredient of *Compound Benzoin Tincture* (page 760). It may be used, like benzoin, to protect fatty substances from rancidity.

Sucrose Octaacetate

α-D-Glucopyranoside, 1,3,4,6-tetra-*O*-acetyl-β-D-fructofuranosyl-, tetraacetate



Sucrose octanecetate [126-14-7] $C_{56}H_{98}O_{19}$ (678.60).

Preparation—Sucrose is subjected to exhaustive acetylation by reaction with acetic anhydride in the presence of a suitable condensing agent such as pyridine.

Description—White, practically odorless powder; intensely bitter taste; hygroscopic; melts not lower than 78°.

Solubility—1 g in 1100 mL of water, 11 mL of alcohol, 0.3 mL of acetone or 0.6 mL of benzene; very soluble in methanol or chloroform; soluble in ether.

Uses—A *denaturant* for alcohol.

Sulfurated Potash

Thiosulfuric acid, dipotassium salt, mixt. with potassium sulfide ($K_2(S_x)$); Liver of Sulfur

Dipotassium thiosulfate mixture with potassium sulfide (K_2S_x) [39365-88-3]; a mixture composed chiefly of potassium polysulfides and potassium thiosulfate. It contains not less than 12.6% of S (sulfur) in combination as sulfide.

Preparation—By thoroughly mixing 1 part of sublimed sulfur with 2 parts of potassium carbonate and gradually heating the mixture in a covered iron crucible until the mass ceases to swell and is melted completely. It then is poured on a stone or glass slab and, when cold, broken into pieces and preserved in tightly closed bottles. When the heat is regulated properly during its production, the reaction is represented approximately by

$$3K_2CO_3 + 8S \rightarrow 2K_2S_3 + K_2S_2O_3 + 3CO_2$$

As this product rapidly deteriorates on exposure to moisture, oxygen and carbon dioxide, it is important that it be prepared recently to produce satisfactory preparations.

Description—Irregular pieces, liver-brown when freshly prepared, changing to a greenish yellow; decomposes upon exposure to air; an odor of hydrogen sulfide and a bitter, acrid, alkaline taste; even weak acids cause the liberation of H_2S from sulfurated potash; 1 in 10 solution light brown in color and alkaline to litmus.

Solubility—1 g in about 2 mL of water, usually leaving a slight residue; alcohol dissolves only the sulfides.

Uses—Extensively in dermatological practice, especially in the official *White Lotion* or *Lotio Alba* (page 762). The equation for the reaction of the potassium trisulfide in preparing the lotion is

$$ZnSO_4 + K_2S_3 \rightarrow ZnS + 2S + K_2SO_4$$

The mixture of insoluble zinc sulfide and sulfur gives the lotion its creamy white appearance.

Talc

Talcum; Purified Talc; French Chalk; Soapstone; Steatite

A native, hydrous magnesium silicate, sometimes containing a small proportion of aluminum silicate.

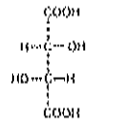
Occurrence and Preparation—The native form, called *soapstone* or *French chalk*, is found in various parts of the world. An excellent quality is obtained from deposits in North Carolina. Deposits of a high grade, conforming to the USP requirements, also are found in Manchuria. The native form usually is accompanied by variable amounts of mineral substances. These are separated from it by mechanical means, such as flotation or elutriation. It then is powdered finely, treated with boiling dilute HCl, washed well and dried.

Description—Very fine, white, or grayish white crystalline powder; unctuous to the touch, adhering readily to the skin, and free from gritti-ness.

Uses—Officially, as a dusting powder and pharmaceutic aid; in both categories it has many specific uses. Its medicinal use as a dusting powder depends on its desiccant and lubricant effects. When perfumed, and sometimes medicated, it is used extensively for toilet purposes under the name *talcum powder;* for such use it should be in the form of an impalpable powder. When used as a filtration medium for clarifying liquids a coarser powder is preferred to minimize passage through the pores of the filter paper; for this purpose it may be used for all classes of preparations with no danger of adsorption or retention of active principles. It is used as a lubricant in the manufacture of tablets, and as a dusting powder when making handmade suppositories. Although it is used as a lubricant for putting on and removing rubber gloves, it should not be used on surgical gloves because even small amounts deposited in organs or healing wounds may cause granuloma formation.

Tartaric Acid

Butanedioic acid, 2,3-dihydroxy-, Butanedioic acid, 2,3-dihydroxy-, [*R*-(*R**,*R**)]



L-(+)-Tartaric acid [87-69-4] $C_4H_6O_6$ (150.09).

Preparation—From *argol*, the crude cream of tartar (potassium bitartrate) deposited on the sides of wine casks during the fermentation of grapes, by conversion to calcium tartrate which is hydrolyzed to tartaric acid and calcium sulfate.

Description—Large, colorless or translucent crystals, or a white granular to fine crystalline powder; odorless; acid taste; stable in the air; solutions acid to litmus; dextrorotatory.

Solubility—1 g in 0.8 mL of water, 0.5 mL of boiling water, 3 mL of alcohol or 250 mL of ether; freely soluble in methanol.

Uses—Chiefly, as the acid ingredient of preparations in which it is neutralized by a bicarbonate, as in effervescent salts, and the free acid is completely absent or present only in small amounts in the finished product. It also is used as a buffering agent.

Trichloromonofluoromethane

Methane, trichlorofluoro-,

$CFCl_3$

Trichlorofluoromethane [75-69-4] CCl_3F (137.37).

Preparation—Carbon tetrachloride is reacted with antimony trifluoride in the presence of a small quantity of antimony pentachloride. The reaction produces a mixture of CCl_3F and CCl_2F_2 which is readily separable by fractional distillation.

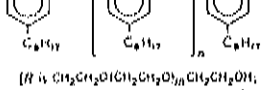
Description—Clear, colorless gas; faint, ethereal odor; vapor pressure at 25° is about 796 torr; boils about 24°.

Solubility—Practically insoluble in water; soluble in alcohol, ether or other organic solvents.

Uses—A *propellant* (No 11, see page 1696).

Tyloxapol

Phenol, 4-(1,1,3,3-tetramethylbutyl)-, polymer with formaldehyde and oxirane; (*Various Mfrs*)



[*R* is $CH_2CH_2O(CH_2CH_2O)_mCH_2CH_2OH$; m is 6 to 9, n is not more than 5]

p-(1,1,3,3-Tetramethylbutyl)phenol polymer with ethylene oxide and formaldehyde [25301-02-4].

Preparation—*p*-(1,1,3,3-Tetramethylbutyl)phenol and formaldehyde are condensed by heating in the presence of an acidic catalyst and the polymeric phenol thus obtained is reacted with ethylene oxide at elevated temperature under pressure in the presence of NaOH. US Pat. 2,454,541.

Description—Amber, viscous liquid; may show a slight turbidity; slight aromatic odor; specific gravity about 1.072; stable at sterilization temperature and in the presence of acids, bases and salts; oxidized by metals; pH (5% aqueous solution) 4 to 7.

Solubility—Slowly but freely soluble in water; soluble in many organic solvents, including acetic acid, benzene, carbon tetrachloride, carbon disulfide, chloroform or toluene.

Uses—A nonionic detergent that depresses both surface tension and interfacial tension. It is a component of Alevaire (*Sterling*) and Enuclene (*Alcon*). It also is used in contact-lens-cleaner formulations.

Zinc-Eugenol Cement

Zinc Compounds and Eugenol Cement NF XI

The Powder

Zinc Acetate	0.5 g
Zinc Stearate	1 g
Zinc Oxide	70 g
Rosin	28.5 g

Powder the rosin and incorporate it with about an equal weight of zinc oxide until thoroughly mixed. Sift the mixture on a sieve of not less than 100-mesh. Regrind the material which does not pass through the sieve with more of the zinc oxide and sift again; repeat the process until all of the material readily passes through the sieve. Thoroughly mix the zinc stearate and zinc acetate with a portion of the zinc oxide and pass through a 100-mesh sieve. Thoroughly mix the two mixtures with the remainder of the zinc oxide.

The Liquid

Eugenol	85 mL
Cottonseed Oil	15 mL

The Cement

To prepare the cement, mix 10 parts of the powder with 1 part of the liquid to a thick paste immediately before use. *Note:* The amount of liquid may be varied to give any desired consistency.

Description—*Powder:* Yellowish white to white in color; *Liquid:* Thin and colorless to weak yellow, having a strong aromatic odor of clove and a pungent, spicy taste; affected by light; specific gravity 1.043 to 1.048; refractive index 1.528 to 1.531 at 20°.

Solubility—*Liquid:* miscible with alcohol, chloroform or ether; only slightly soluble in water.

Uses—In general dental practice as a *dental protective*, ie, as a pulp capping or a *temporary filling*.

Iso-Alcoholic Elixir

Iso-Elixir

Low-Alcoholic Elixir
High-Alcoholic Elixir ..... of each a calculated volume

Mix the ingredients.

Low-Alcoholic Elixir

Compound Orange Spirit	10 mL
Alcohol	100 mL
Glycerin	200 mL
Sucrose	320 g
Purified Water, a sufficient quantity, To make	1000 mL

Alcohol Content—8 to 10%.

High-Alcoholic Elixir

Compound Orange Spirit	4 mL
Saccharin	3 g
Glycerin	200 mL
Alcohol, a sufficient quantity, To make	1000 mL

Alcohol Content—73 to 78%.

Uses—Intended as a general *vehicle* for various medicaments that require solvents of different alcohol strengths. When it is specified in a prescription, the proportion of its two ingredients to be used is that which will produce a solution of the required alcohol strength.

The alcohol strength of the elixir to be used with a single liquid galenical in a prescription is approximately the same as that of the galenical. When galenicals of different alcohol strengths are used in the same prescription, the elixir to be used is to be of such alcohol strength as to secure the best solution possible. This generally will be found to be the average of the alcohol strengths of the several ingredients.

For nonextractive substances, the lowest alcohol strength of the elixir that will yield a perfect solution should be chosen.

Other Miscellaneous Pharmaceutical Necessities

Bucrylate [Propenoic acid, 2-cyano-, 2-methylpropyl ester; Isobutyl 2-cyanoacrylate [1069-56-2] $C_8H_{11}NO_2$ (153.18); (*Ethicon*)]—*Preparation:* One method reacts isobutyl 2-chloroacrylate with sodium cyanide. *Uses:* Surgical aid (tissue adhesive).

Ceresin [Ozokerite; Earth Wax; Cerasin; Mineral Wax; Fossil Wax]—A hard, white odorless solid resembling spermaceti when purified, occurring naturally in deposits in the Carpathian Mountains, especially in Galicia. It is a mixture of natural complex paraffin hydrocarbons. Melts between 61 and 78°; specific gravity 0.91 to 0.92; stable toward oxidizing agents. Soluble in 30% alcohol, benzene, chloroform, petroleum, benzin or hot oils. *Uses:* Substitute for beeswax; in dentistry, for impression waxes.

Ethylenediamine Hydrate BP, PhI [$H_2NCH_2CH_2NH_2.H_2O$]—Clear, colorless or slightly yellow liquid with an ammoniacal odor and characteristic alkaline taste; solidifies on cooling to a crystalline mass (mp 10°); boils 118 to 119°; specific gravity about 0.96; hygroscopic and absorbs CO_2 from the air; aqueous solutions alkaline to litmus. Miscible with water or alcohol; soluble in 130 parts of chloroform; slightly soluble in benzene and ether. *Uses:* In the manufacture of aminophylline and in the preparation of aminophylline injections.

Ferric Oxide, Red—Contains not less than 90% Fe_2O_3. It is made by heating native ferric oxide or hydroxide at a temperature which will yield a product of the desired color. The color is governed by the temperature and time of heating, the presence and kind of other metals and the particle size of the oxide. A dark-colored oxide is favored by prolonged heating at high temperature and the presence of manganese. A light-colored oxide is favored by the presence of aluminum and by finer particle size. *Uses:* Imparting color to mercadamine and cosmetics.

Ferric Oxide, Yellow—Contains not less than 97.5% Fe_2O_3. It is prepared by heating ferrous hydroxide or ferrous carbonate in air at a low temperature. *Uses:* As for *Red Ferric Oxide* (above).

Honey NF XII [Mel; Clarified Honey; Strained Honey] is the saccharine secretion deposited in the honeycomb by the bee, *Apis mellifera* Linné (Fam *Apidae*). It must be free from foreign substances such as parts of insects, leaves, etc, but may contain pollen grains. *History:* Honey is one of the oldest of food and medicinal products. During the 16th and 17th centuries it was recommended as a cure for almost everything. *Constituents:* *Invert sugar* (62 to 83%), *sucrose* (0 to 8%) and *dextrin* (0.26 to 7%). *Description:* Thick, syrupy liquid of a light yellowish to reddish brown color; translucent when fresh, but frequently becomes opaque and granular through crystallization of dextrose; characteristic odor and a sweet, faintly acrid taste. *Uses:* A sweetening agent and pharmaceutic necessity.

Hydriodic Acid, Diluted—Contains, in each 100 mL 9.5 to 10.5 g of HI (127.91), and 600 mg to 1 g of HPH_2O_2 (66.00). The latter is added to prevent the formation of free iodine. *Caution: It must not be dispensed or used in the preparation of other products if it contains free iodine. Preparation:* On a large scale, by the interaction of iodine and hydrogen sulfide. *Description* and *Solubility:* Colorless or not more than pale-yellow, odorless liquid; specific gravity about 1.1. Miscible with water or alcohol. *Uses:* In *Hydriodic Acid Syrup* (page 1302). The latter has been used as an expectorant. It also is used in the manufacture of inorganic iodides and disinfectants. The 57% acid also is used for analytical purposes, such as methoxyl determinations.

Lime [Calx; Calcium Oxide; Quicklime; Burnt Lime; Calx Usta; CaO (56.08)]—*Preparation:* By calcining *limestone* (a native calcium carbonate) in kilns with strong heat. *Description and Solubility:* Hard, white or grayish white masses or granules, or a white or grayish white powder; odorless; solution strongly alkaline. 1 g soluble in about 840 mL of water or 1740 mL of boiling water; soluble in glycerin or syrup; insoluble in alcohol. *Uses:* In making mortar, whitewash and various chemicals and products. It is an ingredient in *Sulfurated Lime Solution* (RPS-16, page 1187). In the USP, calcium hydroxide has replaced it, as it is more stable and more readily available of a quality suitable for medicinal use than that usually obtainable. Unless protected from air, it soon becomes unfit for use, due to the action of carbon dioxide and moisture in the air. See *Calcium Hydroxide* (page 1319).

Peach Oil—An oil resembling almond oil obtained from *Persica vulgaris* (Fam *Rosaceae*). See *Persic Oil* (page 1323).

Polacrilin Potassium [Methacrylic acid polymer with divinylbenzene, potassium salt [39394-76-5]; Amberlite IRP-88 (*Rohm & Haas*)]—Prepared by polymerizing methacrylic acid with divinylbenzene and the

resulting resin is neutralized with KOH. Dry, buff-colored, odorless, tasteless, free-flowing powder; stable in light, air and heat; insoluble in water. *Uses:* *Pharmaceutic aid* (tablet disintegrant).

Poloxalene [Glycols, polymers, polyethylene-polypropylene [9003-11-6]; Bloat Guard (*SmithKline*)]—Polypropylene glycol is reacted with ethylene oxide. *Uses.* *Pharmaceutic aid* (surfactant).

Raspberry Juice—The liquid expressed from the fresh ripe fruit of *Rubus idaeus* Linné or of *Rubus strigosus* Michaux (Fam *Rosaceae*); contains not less than 1.5% of acids calculated as citric acid. *Preparation:* Express the juice from the washed, well-drained, fresh, ripe red raspberries. Dissolve 0.1% of benzoic acid in the expressed juice and allow it to stand at room temperature (possibly for several days) until a small portion of the filtered juice produces a clear solution when mixed with ½ of its volume of alcohol, the solution remaining clear for not less than 30 min. Strain the juice from the mixture or filter it, if necessary. *Description.* Clear liquid with an aromatic, characteristic odor and a characteristic, sour taste; freshly prepared juice red to reddish orange; affected by light; specific gravity 1.025 to 1.045; pH 2.7 to 3.8; refractive index not less than 1.3445. *Uses* In the preparation of *Raspberry Syrup* (page 1302), a *flavored vehicle.*

Sarsaparilla—The dried root of *Smilax aristolochiaefolia* Miller, known in commerce as Mexican Sarsaparilla; or of *Smilax regelii* Killip et Morton, known in commerce as Honduras Sarsaparilla; or of *Smilax febrifuga* Kunth, known in commerce as an Ecuadorian Sarsaparilla; or of undetermined species of *Smilax* Linné, variously known in commerce as Ecuadorian and Central American Sarsaparilla (Fam *Liliaceae*).—Contains glycosides of the saponin group, *sarsasaponin* (*parillin*) and *smilasaponin* (*smilacin*), which are related structurally to the digitalis glycosides, and possess the steroid nucleus. When hydrolyzed with dilute acids, they split into sugars and the corresponding sapogenin. Sarsasaponin yields *sarsasapogenin* (*parigenin*) plus one rhamnose and two glucose molecules, and smilacin yields *smilagenin* plus sugar molecules. Starch, resin, coloring matter and volatile oils also are present. This drug was first used in Europe in the 16th century as a much-vaunted remedy for syphilis. The origin of the name is in doubt. *Uses.* Being without pharmacological actions, it is not employed in modern therapeutics, although the laity is inclined to attribute certain therapeutic virtues to its use.

Sodium Glutamate [Sodium Acid Glutamate [142-47-2] $HOOCCH(NH_2)CH_2CH_2COONa$]—White or nearly white, crystalline powder. Very soluble in water; sparingly soluble in alcohol. *Uses* Imparts a meat flavor to foods.

Sodium Thioglycollate [Sodium Mercaptoacetate; $HSCH_2COONa$]—Hygroscopic crystals which discolor on exposure to air or iron. Freely soluble in water; slightly soluble in alcohol. *Uses* Reducing agent in Fluid Thioglycollate Medium for sterility testing.

Suet, Prepared [Mutton Suet]—Internal fat of the abdomen of the sheep, *Ovis aries* (Fam *Bovidae*), purified by melting and straining. White, solid fat with a slight, characteristic odor and taste when fresh; melts between 45° and 50° and congeals between 37° and 40°; must be preserved in a cool place in tight containers. *Uses* In ointments and cerates.

Urea [Carbamide [57-13-6] $CO(NH_2)_2$(60.06)]—A product of protein metabolism; prepared by hydrolysis of cyanamide or from carbon dioxide by ammonolysis. Colorless to white crystals or white, crystalline powder; almost odorless but may develop a slight odor of ammonia in presence of moisture; melts 132 to 135°. 1 g dissolves in 1.5 mL of water or 10 mL of alcohol; practically insoluble in chloroform or ether. *Uses* A protein denaturant that promotes hydration of keratin and mild keratolysis in dry and hyperkeratotic skin. It is used in 2 to 20% concentrations in various dry-skin creams.

CHAPTER 69

Pharmacological Aspects of Substance Abuse

Frederick J Goldstein, PhD
Professor of Pharmacology

G Victor Rossi, PhD
Vice President of Academic Affairs and Professor of Pharmacology

Philadelphia College of Pharmacy and Science
Philadelphia, PA 19104

Substance abuse continues to be a major problem within the US and will remain so in the 21st century. Although recent data indicate a decline in the use of certain illicit drugs by high-school seniors (Table I), there has been no change in the percent who drink alcohol. Unfortunately, among the many American adults currently dependent upon alcohol and other psychoactive chemicals are members of the health profession, eg, pharmacists, nurses and physicians.

Substance abuse may originate with the *physician*, the *patient* seeking medical treatment or with the *adolescent* drug experimenter. Physician-generated misuse may result when there is insuffcient concern or time to evaluate the patient adequately as a candidate for psychoactive drug therapy. Treatment is all too often directed toward the alleviation of symptoms without a concerted effort to identify possible deep-seated causes and respond to the emotional as well as the medical needs of the patient. Overprescribing of mood-altering drugs involves potential harm not just to the individual but to society at large. While physician-generated drug misuse represents a relatively small percentage of the overall problem, it is especially regrettable that any negative contribution arises from the actions or inactions of health professionals.

Patient-originated abuse encompasses a larger aspect and persists despite significant efforts by the majority of physicians and pharmacists to restrict the dispensing of psychoactive agents. Some patients will visit several physicians, obtain a number of prescriptions for barbiturates, tranquilizers, stimulants and/or narcotics and present each prescription to a different pharmacy. Thus, the patient may accumulate substantial quantities of controlled substances either for personal use or for resale. Attempts to thwart such patterns of drug acquisition have, thus far, been unsuccessful.

Peer pressure, alienation, hedonism, mass-media advertising, affluence and boredom are among the factors most frequently cited as those leading to the misuse of drugs by adolescents. The consumption of alcoholic beverages, cigarette smoking and the liberal use of sedatives, tranquilizers and central nervous system (CNS) stimulants by adults, particularly family members, foster the development of a cavalier attitude toward drugs, and increase the likelihood of drugtaking among adolescents.

Three basic stages of adolescent drug usage have been defined as the initial experimental phase, periodic recreational phase and compulsive (chronic) pattern. That many young people resist involvement with drugs or do not progress to chronic or serious patterns of abuse emphasizes the importance of personality traits in the genesis of drug dependency. Persons of any age who have a low frustration tolerance, cannot cope with the daily pressures of life, require instantaneous gratification or who have unfulfilled dependency needs and serious problems of socialization may come to rely on drug use in order to escape, albeit temporarily, from a psychological environment which is bleak, joyless and/or filled with anxiety.

As stated, many factors are involved in the process by which an individual ultimately selects the pharmacological route of escape from stress. Recent studies indicate that a small percentage of the population may have a genetic predisposition for developing an addiction to at least one drug—alcohol. However, it is quite clear that some potential addicts can resist entering this pathway if they become aware of the toxicological consequences of drug abuse. Many school, religious and community organizations have, in fact, made substantial efforts to present educational programs devoted to acute and chronic toxicities produced by psychoactive substances. Pharmacists should expand their participation in these programs; in this regard, the following information can be of assistance.

Central Nervous System Depressants

Opioids (Narcotics)

Heroin is the opioid most often abused. The preference for heroin is not based on its unique euphoric properties but is largely a matter of economics; heroin is the most potent of the opioids, thus providing maximum profit per kilogram to those engaged in illicit traffic.

Early in the course of heroin use, intravenous injection is followed quickly by a sense of exquisite visceral pleasure which is similar to sexual orgasm (the *rush*), an enveloping feeling of contentment and the receding of internal conflicts. Taken orally, heroin also produces relaxation, euophoria and indifference to pain and stress but not the "rush." In the susceptible individual, the intense desire to recapture this drug experience contributes to the establishment of an emotional or psychic dependency.

With frequently repeated administration, the individual becomes progressively less responsive to the drug; thus,

Table I—National Survey of Lifetime Use[a] of Drugs by High-School Seniors[b]

	1985	1987	Change from 1985
Alcohol	92.2%	92.2%	0.0%
Barbiturates	9.2%	7.4%	−1.8%
Cocaine	17.3%	15.2%	−2.1%
Marijuana	54.2%	50.2%	−4.0%
Methaqualone	6.7%	4.0%	−2.7%

[a] Percent who ever used.
[b] *National Institute on Drug Abuse Notes*, Summer 1988.

everincreasing doses are sought in an attempt to duplicate the characteristic effects. Chronic suppression of central nervous system function results in a dependent state in which the drug must be taken on a regular basis to maintain a reasonable semblance of well-being and equilibrium and to prevent the anguish of the abstinence syndrome. Thus, opioid addicts soon find themselves taking heroin not for the pleasurable effects but primarily to prevent withdrawal.

Tolerance to opioids does not develop uniformly. For example, addicts experience, during chronic use, lessened respiratory depressant, analgesic, sedative, emetic and euphoric effects. Some may show decreased miosis while most suffer chronically from the constipating effects of the drug. Drug tolerance always is relative, never absolute; a dose always exists that is capable of causing death from respiratory paralysis, and overdosage is a common cause of fatalities among opioid addicts. Although death associated with heroin use has been attributed routinely to overdosage, other factors sometimes may be involved.

Quinine frequently is employed by "dealers" to dilute pure heroin because, like the opioid, it is bitter and produces vasodilation simulating the *rush*. Thus, addicts cannot detect adulteration readily and may unknowingly inject themselves with large quantities of quinine, which may produce significant myocardial depression. Codeine, while significantly less potent than heroin, also can produce death from overdosage.

Withdrawal symptoms usually reach maximum intensity 36 to 72 hr after the last dose of heroin and subside gradually within 7 to 10 days. The severity of the abstinence syndrome is determined largely by the degree of acquired physical dependence and the rate of elimination of the drug.

The signs and symptoms of opioid withdrawal include yawning, sneezing, lacrimation, restlessness, anxiety, insomnia, nausea, vomiting, gastrointestinal cramps and diarrhea, sweating, gooseflesh, generalized body aches, fever, tremors, muscle spasms and jerking movements. Excessive perspiration, vomiting and diarrhea combined with diminished food and fluid intake may result in dehydration, acid-base disturbances and ketosis. Occasionally, cardiovascular collapse occurs.

Withdrawal symptoms can be suppressed either by administering the drug of dependence or another narcotic. If an opioid, such as methadone, is given initially in a stabilizing amount and then the dosage reduced gradually, the intensity of the abstinence syndrome may be lessened appreciably.

The opioid addict is subject to risks arising out of indifference to minimal nutritional and hygienic requirements with a consequent high incidence of viral hepatitis, bacterial endocarditis, tetanus, pulmonary infection, pulmonary edema and thrombophlebitis.

The use of nonsterile injection equipment and intravascular introduction of cotton fibers and adulterants, such as lactose and talc, all contribute to the development of local and systemic infectious disorders and pulmonary granulomatosis. Hyperamylasemia often is observed during the acute phase of heroin-induced pulmonary disturbances. Increased serum immunoglobulin levels are encountered commonly in addicts. Although the clinical consequences of this finding are understood incompletely, serological tests for syphilis are false-positive in a significant proportion of such individuals.

Noninfectious complications of opioid addiction include transverse myelitis, rhabdomyolysis with cardiac involvement and myoglobinuria and Horner's syndrome. Quinine contained in *street* heroin preparations produces amblyopia and thrombocytopenia. An aqueous mixture consisting of crushed tablets of pentazocine (Talwin) and tripelennamine (Pyribenzamine), with the street name of "T's and Blues', has been used intravenously by addicts; the effects are reported to be similar to the heroin *rush*.

Toxic reactions can be serious and include tonic-clonic seizures and acute respiratory distress with hypoxia. The latter effects apparently result from deposition of insoluble ingredients of this mixture, eg, talc, in lung tissue thus causing pulmonary granulomas.

MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine), an extremely toxic by-product of illicit meperidine synthesis, destroys certain types of brain tissue (nigrostriatal) after only a few doses; this produces Parkinson's disease in the abuser which, like the degenerative clinical disease occurring in geriatric patients, is irreversible.

Women who persist in the use of heroin during pregnancy give birth to opioid-dependent offspring. The signs of withdrawal in the newborn appear within several hours to several days and include high-pitched crying, sleeplessness, irritability, tremor, vomiting and diarrhea; the latter may result in severe dehydration. Narcotic-dependent infants are born smaller and exhibit an uncoordinated and ineffectual sucking reflex, which reduces nutritive consumption. Phenobarbital, diazepam, paregoric or chlorpromazine have been used to alleviate narcotic withdrawal in neonates.

The approaches to treatment of the adult addict involve medical as well as psychiatric and social aspects. A basic obstacle in any approach to the treatment of opioid addiction is the characteristic high rate of recidivism.

Methadone maintenance, currently one of the most widely employed techniques in the management of opioid addiction, involves stabilizing the patient on a regular daily oral dose of methadone, preferably in conjunction with supportive psychological or psychiatric counseling. In this context, the maintenance drug does not provide true pharmacological blockade; rather, regular administration results in the development of tolerance to methadone and cross-tolerance to heroin. Thus, the addict will not experience the heroin-induced "rush" and euphoria unless doses substantially higher than usual are injected.

Theoretically, when unburdened by these factors which motivate addiction, methadone may be withdrawn gradually. However, many former narcotic abusers cannot maintain a drug-free state and either reestablish their addiction to heroin or request continued metadone therapy. In contrast, some addicts refuse to enter a methadone maintenance program. The reasons for this decision include

The claim by some narcotic abusers that methadone is just another type of drug dependence and one which is more difficult to surrender than heroin use—in fact, methadone withdrawal can be more intense and painful than heroin detoxification.

Methadone significantly impairs human reproductive capacity by decreasing both ejaculate volume and sperm motility (heroin produces a lesser effect upon fertility).

Family members may be endangered—a number of children have died after ingesting the liquid methadone preparations used by their parents.

An alternative approach, based on the conditioning theory of opioid dependence, employs narcotic antagonists to extinguish drug-seeking behavior by blocking the euphoric effects of heroin. Nalorphine first was suggested for this purpose but its limited duration of action and high incidence of hallucinogenic reactions made its use impractical. Cyclazocine is effective orally and provides blockage for up to 24 hr but, like nalorphine, is an active analgesic and is associated with a variety of disturbing psychotomimetic reactions. Naloxone (*Narcan*), a "pure" opioid antagonist (ie, possesses no agonist properties), produces fewer unpleasant effects but is relatively short-acting.

Naltrexone (*Trexan*), a longer-acting derivative, can block the effects of heroin for approximately 72 hr. The results of clinical trials are disappointing since many addicts

under treatment refuse to take a drug which is devoid of narcotic-like effects.

Clonidine (*Catapres*), an α_2-receptor-agonist, has been used successfully in treating heroin withdrawal reactions; in some cases, it is more efficacious than methadone.

Barbiturates

The clinical use of barbiturates has declined substantially in recent years. The benzodiazepines, while not free of adverse reactions, are safer and have supplanted barbiturates in the treatment of anxiety and insomnia. It is clear that, in general, hypnotics (barbiturates and nonbarbiturates) should not be prescribed for more than a 14 to 28-day period. Beyond this time efficacy decreases (a decline in hypnotic activity may begin after only 7 days of continuous therapy). Pharmacists should monitor these prescriptions very closely, consulting with both the patient and physician in order to insure proper use and prevent dependence problems.

The hazards encountered in the use of barbiturates include occasional unanticipated idiosyncratic or hypersensitivity reactions and accidental overdosage as may occur in young children unaware of the potential danger or in adults during a hypnotic drug-induced semistuporous state of "automatism." For most persons, sleep provides only a temporary respite but, all too frequently, intentional overdosage with easily accessible sleep-inducing drugs provides an avenue of permanent escape from the pressures of reality.

Barbiturates reduce the amount of time spent in the REM (rapid eye movement) phase of sleep. The reduction of REM sleep for a period of several days may cause the individual to become irritable or to evidence disturbances in personality and rationality. When the hypnotic is withdrawn abruptly, there is a rebound increase in the REM phase often associated with nightmares, a feeling of having slept poorly or actual insomnia. "Rebound" REM makes it difficult for the patient to give up the drug and contributes to the development of drug dependency.

The signs and symptoms of barbiturate and alcohol intoxication are strikingly similar. Visual perception, recall, reaction-time coordination and other indexes of psychomotor functioning are affected, the degree of impairment largely depending on the concentration of drug in the brain. Intoxication, either with alcohol or a barbiturate, is characterized by difficulty in thinking, reduction of ego controls, poor judgment, confusion and emotional instability. Neurological impairment and muscular incoordination are major factors in the personal injuries and involvement in vehicular accidents which are common occurrences during the course of intoxication with these drugs. The CNS suppressant effects of alcohol, barbiturates and opiates, such as heroin, are mutually reinforcing; extemporaneous combinations of these depressants may result in unpredictably abrupt and severe incapacitation.

Low doses of barbiturates (as employed for daytime sedation, nighttime sleep induction or the control of epilepsy) are often taken for indefinite periods without eliciting tolerance or physical dependency. These phenomena generally occur only with doses considerably in excess of those customarily employed in medical practice. To illustrate, the usual oral hypnotic dose of pentobarbital sodium or secobarbital sodium is 100 to 200 mg, whereas oral doses of these barbiturates in excess of 400 mg/day (and generally in the range of 600 to 800 mg/day) for approximately 1 month are required to induce clinically significant tolerance and physical dependency. Parenteral (subcutaneous or intravenous) administration of barbiturates may lead to physical dependency at lower dose levels and within a shorter period of time.

The amount of barbiturate that may be consumed by the compulsive abuser varies considerably, but average daily doses of 1 to 1.5 g of short-acting derivatives are not uncommon, and some individuals may use as much as 2.5 g/day over prolonged periods of time.

Withdrawal reactions, which in some cases may be more hazardous than the opioid abstinence syndrome, develop upon abrupt cessation of chronic barbiturate overuse. Mild to moderate withdrawal reactions include anorexia, apprehension, tremulousness, muscular weakness, mental confusion and postural hypotension. A severe barbiturate withdrawal syndrome may involve profound disorientation, delirium and hallucinations and convulsive seizures of an episodic or protracted nature. Most individuals who have ingested eight or more hypnotic doses of a barbiturate per day over an extended period will experience convulsions during withdrawal. In extreme cases the barbiturate abstinence syndrome may terminate in cardiovascular collapse and death. With the longer-acting barbiturates, withdrawal symptoms are slower in onset and less severe than those encountered with the shorter-acting derivatives.

Pharmacological treatment of barbiturate dependency generally is approached by replacement with either pentobarbital or phenobarbital at an initial dose sufficient for stabilization; the dose then is reduced gradually over a period of several days to weeks depending on the individual patient response.

Nonbarbiturate Sedative-Hypnotics

Neurological impairment, psychological and physical dependency, and an abstinence syndrome similar to that associated with barbiturate abuse may result from excessive use of many nonbarbiturate sedative-hypnotic and antianxiety agents, including chloral hydrate, glutethimide, methyprylon, methaqualone, meprobamate, chlordiazepoxide or diazepam.

Methaqualone remains a "street" drug of choice. Although claims have been made that it and other nonbarbiturate hypnotics (eg, chloral hydrate or triclofos) produce little or no effect on REM sleep, other reports challenge this distinction and a final conclusion has not been advanced yet.

Acroparesthesia (tingling and numbness in the extremities) may occur prior to the onset of hypnotic activity, particularly when sleep does not ensue rapidly. This sensation is experienced by many methaqualone abusers and probably contributes to the aphrodisiac effect (similar to the 'Spanish Fly' phenomenon). Increased muscle tone often is evident; it even may be observed while the patient is in a deep coma and may last for several days. Acute toxicity differs from that of the barbiturates in that marked respiratory and cardiovascular depression generally are not seen after large doses of methaqualone.

Psychological dependence and tolerance to methaqualone have been observed, but the results of studies on the development of physical dependence are equivocal. Apparent withdrawal symptoms, such as headache, anorexia, nausea, abdominal cramps and interference with sleep, have been noted in those investigations reporting physical dependency. These relatively minor symptoms may occur during abstinence in the individual who has been taking five hypnotic doses of methaqualone daily for several months.

Severe reactions which may be encountered occasionally during methaqualone withdrawal include convulsions and toxic psychoses. Ingesting alcohol with methaqualone is very dangerous, leading to a serious impairment of judgment and psychomotor coordination. At least one state reports a high death rate from injuries sustained in car accidents where the drivers, passengers and/or pedestrians used this drug combination.

Mandrax, a combination of methaqualone (250 mg) and

diphenhydramine (25 mg), has been abused by addicts in Great Britain, Canada and Australia. The reactions due to overdosage with this drug combination are similar to those of methaqualone but are potentially more severe since diphenhydramine, which possesses central antimuscarinic activity, may produce psychological disturbances, excitation, ataxia and convulsive seizures (diphenhydramine does not influence the absorption or biotransformation of methaqualone).

Meprobamate produces sedation and relaxation comparable to that of the barbiturates, although the clinically effective dose of meprobamate is higher. Cognitive activity may be compromised by chronic oral doses of 800 mg of meprobamate per day, while at daily doses of more than 1600 mg, psychomotor performance may be reduced significantly. Psychic dependence and tolerance occur with prolonged high-dose administration and physical dependence develops after consumption of 3 g or more per day for several weeks. Depending on the dosage and duration of use, meprobamate withdrawal reactions may range from anxiety, insomnia and tremors to hallucinations, convulsions, coma and death.

Chlordiazepoxide, taken in doses of 300 to 600 mg a day for several months, may result in physical dependency resembling that observed with the barbiturates and meprobamate. However, withdrawal symptoms may be delayed for several days after chlordiazepoxide is terminated, due possibly to slow elimination of the drug. Agitation, insomnia, anorexia, depression, psychological disturbances and convulsions are among the reactions which follow the cessation of prolonged administration of high doses of chlordiazepoxide.

Diazepam, the most widely prescribed benzodiazepine derivative, also may induce physical dependence. Patients receiving 15 mg a day for 4 to 6 months, or higher doses (60 to 120 mg) for about 2 months may, upon withdrawal, experience gastric cramps, sweating, agitation, tremors, insomnia, confusion, disorientation, auditory and visual hallucinations, delusions, paranoia and depression.

Serious acute intoxication may occur when benzodiazepines are combined with other depressants, eg, ethanol, narcotics, other sedative-hypnotics, tricyclic antidepressants or antipsychotic agents. Simultaneous ingestion of ethanol and diazepam is particularly dangerous. In addition to the expected additive CNS-depressant effects, in the presence of ethanol, diazepam blood levels are elevated, compared to diazepam taken alone. Some reports suggest the possibility of teratogenicity resulting from administration of meprobamate or certain benzodiazepines during the first trimester; in the interest of caution, the use of these antianxiety agents should be restricted during this critical period of pregnancy.

The medical and pharmaceutical professions bear a grave responsibility in prescribing and dispensing barbiturates, benzodiazepines and pharmacologically related agents. Physicians, pharmacists and nurses often fail to convey adequately to the patient the potential of these drugs for ensnarement in a vicious web of emotional need, often progressing to escalated consumption and, ultimately, the development of a dangerous degree of psychological and physiological dependency. Although only a limited number of drugs were discussed in the above sections, it is important to note that *any* substance causing acute CNS depression is capable of producing psychological and/or physical dependency during chronic use.

The legitimate application of drugs should not be jeopardized by irrational fears arising from situations created by their uncontrolled use. However, it is equally important to recognize that certain drugs, by virtue of their ability to elicit profound changes in mood and feeling, may, in the emotionally predisposed person, lead to a degree of psychic dependency and compulsive use detrimental to the individual and to society.

Alcohol

Although greater publicity usually is accorded to marijuana, hallucinogens and narcotics, alcohol remains the major drug of abuse in the US. Approximately 15% of all US health costs are for the treatment of chronic alcoholism and associated toxicities.

Alcoholic intoxication spans a range of blood-ethanol concentrations from 0.05%, at which level some impairment of judgment occurs, to above 0.40%, associated with profound depression of vital physiological functions. Concentrations in excess of 0.60% usually are fatal.

Although many states regard an individual as being "legally drunk" at levels above 0.10%, controlled studies have demonstrated repeatedly that functional deficits such as impaired adaptation to light, reduced psychomotor performance with prolonged reaction times and generalized deterioration of simulated driving skills are evident at blood-alcohol concentrations well below 0.08%. Thus, individuals with blood-alcohol levels below those required for legal classification as intoxicated may, nevertheless, be dangerous drivers.

Compelling statistics compiled over many years implicate alcohol as a principal contributor to motor vehicle accidents with consequent injuries and fatalities. Public outrage by groups such as *Mothers Against Drunk Drivers* (*MADD*) has been directed recently toward the legislative and judicial systems for their minimal penalization of drunk drivers, particularly the repeat offender. As a result, most states now have passed stricter laws with more severe penalties. All 50 states now require a person to be 21 yr old in order to drink alcohol.

Two-day jail terms for first offenses and quicker suspension of the operator's license now are routine aspects of punishment. However, none of these statutes can restore the lives of innocent children and adults who have been killed by intoxicated drivers. The prevention of alcohol abuse through educational and other methods remains the approach most likely to reduce deaths. Many airline pilots and railroad engineers currently are involved in such programs.

Severe alcoholic intoxication may result in forms of amnesia characterized as "state-dependent learning" or as a "blackout." In the former, an individual can recall what transpired under the influence of alcohol only if again subjected to an intoxicated state. Generally, information acquired while under the effects of alcohol is remembered poorly or not retained in the nondrug condition.

"Blackout" refers to a severe short-term memory deficit; subjects cannot recall what occurred while intoxicated even if they again become inebriated. Assaultive or destructive behavior (eg, suicide, attempted suicide or homicide) associated with drinking frequently takes place during an amnesic state.

Estimates of the number of alcoholics in any society are very imprecise; the number of individuals in the US alone, whose lives are involved inextricably with alcohol, is numbered conservatively in the several millions. The cost in terms of lost productivity, accidents, crimes, self-degradation and the disruption of family, business and social bonds is beyond computation. Chronic abuse leads to debilitating pathological alterations which seriously impair the alcoholic's health and diminish life expectancy; these effects may be summarized as follows:

1. *Mortality*
   The probability of premature death is approximately three times that of the general population, in addition to a greater frequency of fatal accidents and suicides; pathological changes are contributory.
2. *Cardiovascular*
   While several clinical studies show a reduced incidence of heart disease (possibly due to elevation in protective serum high-density lipoproteins) among persons who consume an average of 2 oz or less of alcohol per day, heavy drinkers (more than 2 oz a day) are at greater risk of developing various cardiovascular disorders which include:

a. permanent dilation of peripheral blood vessels around nose and eyes
b. hypertension
c. artherosclerotic heart disease
d. congestive heart failure
e. peripheral vascular disease
f. cerebrovascular disease (eg, stroke)

3. *Neurological*
Observed clinical changes may occur as:
a. cerebellar ataxia (motor incoordination)
b. decreased ability to perform cognitive tasks (eg, verbal and nonverbal tests)
c. polyneuropathy
d. nystagmus
e. Korsakoff psychosis
f. Wernike encephalopathy (may include some or all of above, ie, 3a to 3e)
Cerebral atrophy, documented by computerized axial tomography, can be extensive and has been linked to functional neurological deficits. Of particular interest is a report which suggests the loss of cognitive skills may be related more to consumption of substantial amounts of alcohol per drinking episode than to the frequent use of limited quantities. Partial recovery may occur with total abstinence.

4. *Hepatic*
Degenerative alterations in liver morphology and function appear during chronic alcoholism and develop progressively in the following order (includes sequelae):
a. alcoholic fatty liver (hepatic pain and tenderness)
b. alcoholic hepatitis (nausea, vomiting, anorexia, weight loss, abdominal pain)
c. cirrhosis (jaundice, encephalopathy)
As with alcohol-induced neurological changes, cessation of drinking usually prevents further deterioration.

5. *Gastrointestinal*
Ulcer formation and extensive gastrointestinal bleeding frequently are seen in addition to:
a. esophagitis
b. gastritis
c. intestinal malabsorption (of, for example, fat, folic acid, thiamine, vitamin B_{12})
d. chronic diarrhea
e. steatorrhea

6. *Pancreatic*
Chronic pancreatitis often is observed after approximately 7 yr of heavy alcohol use (usually appears before cirrhosis). Pancreatic failure may produce insulin-dependent diabetes mellitus.

7. *Hematological*
Anemia may be caused by deficiencies of folic acid and/or iron; other disorders are:
a. thrombocytopenia
b. granulocytopenia

8. *Endocrine*
a. diabetes mellitus
b. pseudo-Cushing's syndrome
c. hypogonadism
(1) female: amenorrhea
(2) male: low plasma testosterone levels, impotence, infertility, testicular atrophy

9. *Infection*
a. bacteremia
b. bacterial peritonitis
c. pneumonia
d. tuberculosis

10. *Cancer*
a. esophageal
b. hepatic
c. laryngeal
d. pharyngeal
e. mouth
f. breast (possibly)

Although alcoholic beverages constitute an appreciable source of calories, they provide no vitamins, minerals or proteins. Nutritional deficiencies associated with long-term heavy drinking may constitute major factors in the development of polyneuritis and cirrhosis of the liver. However, evidence suggests that liver damage results from the direct hepatotoxic effect of alcohol and/or its metabolites and that cirrhosis may occur independently of nutritional status.

Alcohol passes readily from the maternal to the fetal circulation, thus frequent consumption of alcohol during pregnancy creates an unnatural intrauterine environment for the developing fetus. Infants born to alcoholic mothers usually are underdeveloped and exhibit a slow growth rate and mental retardation. Current evidence suggests that these effects may be permanent. Cardiovascular aberrations, including systolic murmurs (due to possible ventricular septal defects) and congestive heart failure (resulting from possible atrial septal defects), and craniofacial abnormalities (such as short palpebral fissures and maxillary hypoplasia) have been documented as patterns of malformation in infants born to chronic alcoholic women. This dysmorphic pattern has been classified as the Fetal Alcohol Syndrome (FAS) and is most likely to occur when maternal consumption is equivalent to 90 mL (or more) of absolute alcohol per day.

The chronic ingestion of alcohol results in pharmacodynamic and drug-disposition tolerance. However, the degree of tolerance is not as great as that which occurs with morphine. Physical dependence develops to alcohol, which is similar to that observed with barbiturates and narcotics. The severity of the alcohol-abstinence syndrome can be correlated with the degree of intoxication and its duration. A relatively short period of heavy drinking may be followed by headache, nausea, vomiting, general malaise and slight tremulousness during the "drying-out" period.

Abrupt cessation of alcohol consumption after 1 week or more of intoxication may be associated further with anxiety, insomnia, confusion, tremors and hallucinations. Long periods of intense intoxication may, upon withdrawal, result in delirium tremens, a syndrome characterized by increased autonomic activity (eg, fever, sweating and tachycardia), agitation, disorientation, severe tremors or convulsive seizures and frightening hallucinations, usually of a visual form.

Hereditary predisposition, endocrine abnormalities, psychological defects, susceptible personality structure and sociocultural and economic impacts are among the many factors that have been considered as interacting in the causation of alcohol addiction. Because of the many conflicting hypotheses on the etiology of alcoholism, there is no standard approach to therapy. There is a general agreement, however, that a prerequisite for successful therapy is *total* abstinence from alcohol and, for all practical purposes, this represents the only viable solution for the individual alcoholic.

Efforts to correct the drinking habit almost invariably fail if the patient attempts merely to reduce his consumption of alcohol. Indeed, the failure of the alcoholic to accept the realization that he is incapable of drinking in moderation is regarded as a primary obstacle to the ultimate resolution of the problem.

Some alcoholics stop drinking of their own volition, others are able to discontinue the habit with the aid of professional or peer-group counseling and still others continue to relapse despite repeated and intensive rehabilitative efforts. Therapeutic measures employed, with varying degrees of success, in the long-term management of the alcoholic patient include participation in supportive social organizations for combating alcoholism (eg, Alcoholics Anonymous), psychiatric therapy and the use of neuroleptic or antianxiety agents, although the latter may result in substitution of one form of drug dependence for another. The unpleasant interaction between alcohol and disulfiram may be used both as a deconditioning device and as a deterrent.

Volatile Hydrocarbons

Volatile hydrocarbons (eg, glue, carbon tetrachloride, gasoline, nail polish remover, lighter fluid, paint, lacquer, varnish thinner—even those solvents found in typewriter correction fluid and adhesive tape remover!) are abused most frequently by young individuals between 10 and 15 yr of age. These liquids usually are deposited in a handkerchief, rag or bag which is then placed over the nose and mouth and the vapors inhaled, a process known as "huffing." Initial exhilaration and CNS excitation may occur with blurring of vi-

sion, ringing in the ears, slurred speech and staggering gait. These effects generally last from 30 to 45 min after inhalation. Depending upon the quantity of vapor inhaled, drowsiness, stupor and unconsciousness may result.

Occasionally, volatile hydrocarbon abuse precipitates psychotic behavior, but susceptible individuals are apparently those who manifest personality disturbances antecedent to drug use. Amnesia often follows recovery. In extreme cases of intoxication, death due to respiratory paralysis may occur.

Psychological dependence can develop and, although physical dependency does not, this latter situation probably is attributable primarily to the limited duration of volatile hydrocarbon use, rather than to the pharmacological properties of these chemicals. If volatile hydrocarbons were abused frequently and for a sufficiently long period, physical dependency might be established, as is the case with other potent CNS depressants, eg, barbiturates and narcotics.

Physical signs associated with the use of volatile hydrocarbons include characteristic odors, irritation of mucous membranes and elevated pulse rate. Chronic abuse may produce damage to the kidneys, liver, heart and brain. In glue sniffers with sickle-cell disease, severe anemia has been observed, possibly as a result of bone-marrow depression. Chromosome damage in glue sniffers has been reported but this adverse reaction remains to be established definitely.

The inhalation of butyl nitrites (primarily isobutyl nitrite) produces pharmacological effects similar to amyl nitrite. Butyl nitrites are found in room deodorizers which contain one or more isomers (eg, *n*-butyl, *sec*-butyl, isobutyl, *tert*-butyl). Euphoria, the most desired immediate effect, often is accompanied by dizziness, fainting, cutaneous flushing, headache and hypotension, all of which are due to significant peripheral vasodilation. Subsequent effects include dermal irritation leading to lesions on the lips, nose, penis and scrotum. Since nitrites may be carcinogenic, chronic inhalation may produce cancer. Many homosexual men have used nitrite inhalants, which may promote Kaposi's sarcoma, commonly found on the nose tips and oral mucosa of such individuals who ultimately contract AIDS.

Aerosol Propellants

More than 2 billion aerosol spray cans are produced each year for such diverse applications as household cleaners, furniture waxes, insecticides, hair sprays, antitussives, paints, antisticking coatings for cookware, deodorizers, disinfectants and cocktail-glass chillers. Many of these aerosols also are widely abused by youthful drug experimenters, primarily teenagers.

The effects which result from "huffing" aerosols generally are similar to those described for volatile hydrocarbons. Reports in the medical literature have described several cases of collapse and death of young persons within a very short time after deliberate inhalation of the contents of various aerosol containers. This phenomenon has been designated "sudden sniffing death" (SSD). The appellation implies a greater degree of specificity, however, than may be warranted. The mechanisms involved in SSD have yet to be elucidated. Autopsy findings have been negative in that no anatomical cause of death has been established. Suffocation, frozen vocal cords and respiratory failure may accompany SSD but do not appear to be the primary factors, since death occurs so rapidly.

Considerable attention has been directed to the fluoroalkane propellant gases (most often Freons) as possible causative agents of SSD. The data provided by some experimental animal studies suggest that the fluoroalkanes are capable of producing direct myocardial depression, bradycardia, atrioventricular block and ventricular dysrhythmias. Other studies conducted with these chemicals, however, have not revealed significant direct cardiotoxicity. Fluoroalkane propellants and volatile hydrocarbon solvents also may have an indirect action on the heart, ie, sensitization of myocardial tissue to the arrhythmogenic effect of the catecholamines. Thus, in individuals exposed to inordinate concentrations of these materials, endogenous epinephrine released during severe stress or physical activity might be expected to produce a markedly deleterious effect on cardiac function. Hypercapnia, as would result from rebreathing the air in a small, closed environment (eg, bag sniffing), may potentiate further the cardiotoxicity of catecholamine, fluoroalkane or volatile hydrocarbon combinations.

Asthmatic patients have been found dead surrounded by one or more bronchodilator aerosol containers, the contents of which have been expended. Investigations into the nature of such fatalities indicate that a severe asthmatic attack itself may be the major cause of death. However, it also has been suggested that fluoroalkane propellants combined with epinephrine or isoproterenol may produce lethal cardiac arrhythmias if the recommended dose of inhalant is exceeded.

Isolated reports have linked the appearance of sarcoid-like lesions in the lungs and premalignant pulmonary lesions to the increased use of aerosol preparations. However, the validity of the presumed association remains to be confirmed.

Deaths related to aerosol propellant abuse have declined during the past few years. This trend apparently is due to elimination of Freons from spray cans in order to prevent environmental damage (eg, destruction of ozone layer in upper atmosphere).

Nitrous Oxide

Inhaling nitrous oxide for nonmedical purposes, ie, to induce a "high," remains a current national problem which is not confined to teenagers. Students at both the college-undergraduate and health-professional level, as well as licensed practitioners, are known to be among the abusers. Supplies of nitrous oxide have been obtained through the theft of large cylinders (eg, as used in hospitals) or the purchase of whipped-cream cartridges which contain approximately 3 L of nitrous oxide.

Acute, uncontrolled exposure can be lethal by promoting unconsciousness in the user who then collapses into a body position which could be suffocating. At least one death has occurred in this manner. Other fatalities are known and the Drug Enforcement Administration estimates that nitrous oxide-related deaths are underreported.

Chronic toxicity develops not only in abusers but also in health professionals who employ nitrous oxide for legitimate purposes. An extensive survey of dentists and dental assistants found that when exposure was "heavy," ie, more than 3000 hr over a 10-yr period (6 hr per week), the number of reported adverse effects was four times greater than those experiencing "light" exposure, ie, less than 3000 hr per 10 yr. The initial signs and symptoms of nerve damage occur as numbness and paresthesias (unusual feelings in limbs described as burning and/or tingling). Later, muscle weakness and gait disturbances may develop. In abusers, this polyneuropathy could become permanent. Other effects of prolonged use which are firmly linked less include headaches, nephrotoxicity, hepatotoxicity, neoplastic disease, spontaneous abortions (higher than normal rate) and teratogenicity.

Marijuana (Marihuana)

Marijuana is obtained from one of man's oldest cultivated plants, *Cannabis sativa*. The biologically active principles of cannabis are concentrated in the resinous exudate of the flower clusters. Traditionally, the female plants have been

harvested for their high resin yield. Chemical analyses have indicated, however, that the cannabinoid content of the resin does not differ significantly between the male and female plants. The potency of preparations derived from cannabis varies enormously depending on their composition and method of formulation. *Hashish*, the unadulterated resin from the flowering tops of cultivated female plants, is a most potent form.

By legal definition (US Federal Statutes), the term *marijuana* embraces all parts, extracts, derivatives or preparations of cannabis, including the pure resin. However, as usually encountered in the Western hemisphere, marijuana comprises a mixture of the leaves, flowering tops and other structural parts of the cannibis plant, generally dried, chopped and incorporated in a form for smoking.

Although Δ^9-tetrahydrocannabinol (THC) appears to be the major active constituent of marijuana, biological activity may be attributable largely to the 11-hydroxy metabolite. Marijuana cigarettes ordinarily obtainable in the US contain about 1 to 2% THC. Based on an average cigarette weight of approximately 500 mg, the amount of available THC ranges from 5 to 10 mg. Stronger products, ie, those with 3 to 5% THC, are currently available in the American "market."

Depending on its potency, a marijuana cigarette will produce moderate to intense psychopharmacological effects which reach a peak within 15 min and persist for 1 to 4 hr. As compared to smoking, marijuana consumed orally is about ⅓ as potent and the onset of activity is delayed but markedly prolonged.

One of the most consistently demonstrable effects of marijuana in humans is elevation of the pulse rate; the rate may rise by 50% or more above the preexposure level and increases may be sustained for several hours. Within limits, the intensity of this response appears to be related to the amount of drug consumed. Blockade by propranolol implicates β-adrenergic receptor activation in the mechanism of THC-induced tachycardia. However, that the increase in heart rate occurs without a simultaneous increase in left ventricular performance suggests the operation of an antivagal mechanism by THC. Smoking marijuana while taking other drugs known to produce tachycardia, eg, nortriptyline, can result in a very substantial elevation of heart rate.

Blood pressure changes are variable; slight elevations and reductions of systolic and diastolic pressure have been noted. Continuous electrocardiographic monitoring of subjects who smoked cigarettes calibrated to contain 20 mg of THC revealed no ECG alterations that could be attributed definitely to marijuana intoxication. In contrast to the increased heart rate observed in humans, THC produces bradycardia in several animal species, eg, the rat, cat or dog.

Reddening of the conjunctiva (conjunctival congestion) is another consistent response to marijuana. That reddening also occurs after oral administration of THC indicates that this is not an artifact produced by irritation from smoke. Despite a belief long associated with marijuana, significant changes in pupillary diameter are not observed. Although marijuana does not elevate the respiratory rate, oral administration may produce airway dilation, probably by direct relaxation of bronchial musculature, for a period of several hours.

Appetite is stimulated in human and subhuman species, but without concurrent alteration of the blood glucose level. Weight gain, which often occurs during prolonged use of marijuana, probably is related more to increased caloric intake than to excessive fluid retention. Disturbances of equilibrium and muscular coordination as well as hyperreflexia during marijuana intoxication have been reported. Other physiological changes noted with marijuana include dryness of the mouth and throat, irritation of the oropharyngeal mucosa, nausea and occasional vomiting, tinnitus and paresthesias.

The marijuana-induced state is characteristically a hypersuggestible state; psychological and perceptual effects are influenced markedly by the mental attitude, mood expectations of the user and the setting and circumstances attending its use. Typically, there is a sense of relaxation, inner contentment, euphoria or even elation; thoughts flow in disconnected fashion in a dream-like state; time and space orientation are impaired; body image is distorted; perception of colors and sounds is altered, usually intensified; laughter comes easily and may be uncontrollable but sometimes mood is subdued or depressed.

The subjective responses to marijuana correlate generally with the onset and duration of tachycardia and conjunctival vascular congestion. EEG changes have been recorded in THC-treated animals, and it has been suggested that the activation of septal areas associated with pleasure and emotion may play a role in certain of the observed psychological alterations.

Short-term memory frequently is impaired and information learned while under the influence of marijuana is recalled effectively only when the individual again is subjected to the drug effect, ie, state-dependent learning. Intense depersonalization, loss of insight, disorganized thinking and speech and grossly distorted perception occur with high doses but true hallucinations rarely are experienced, except at toxic levels. This contrasts with the hallucinogenic drugs (eg, LSD, DMT) which induce organized visual illusions and hallucinations at subtoxic doses.

Performance in psychometric tests is affected variably, depending on the nature of the task, its complexity and the dose of marijuana. Generally, marijuana produces a dose-related psychomotor performance decrement. In tests of driving skills, speedometer errors were increased but braking, signaling or steering responses essentially were unimpaired. There is, however, a significant delay in light adaptation which may seriously impair driving at night. Marijuana prolongs the time needed to regain normal vision after exposure to bright light as, for example, from the headlights of an oncoming automobile. This effect is dose-related and may persist for 2 hr after marijuana use.

That deficiencies in these responses may contribute to automobile accidents is suggested by the finding of measurable blood levels of THC in some motorists involved in traffic violations. In a recent study, subjects with plasma THC levels above 25 to 30 ng/mL failed coordination tests routinely given to drivers to assess the severity of alcohol intoxication. However, the temporal correlation between plasma THC levels and degree of incoordination was not as accurate as with alcohol.

Adverse reactions to marijuana occur relatively infrequently. They have been classified by Weil[1] as follows:

1. Normal population.

   Simple depressive reactions—occur in neophyte users; terminate spontaneously.

   Panic reactions—occur mainly in individuals who have inhibitions regarding use of psychoactive drugs; patient may be anxious, depressed, fearful, withdrawn or agitated but, generally, is panicked due to physiological and/or psychological effects which are misinterpreted as life-threatening.

   Toxic psychoses—serious, temporary disturbances of normal brain activity; patients are disoriented and frequently experience hallucinations.

2. Persons who previously have taken hallucinogenic drugs.

   Precipitation of "flashbacks".—marijuana may induce recurrences of a "trip" which developed originally from previous consumption of a hallucinogenic drug.

   Precipitation of delayed psychotic reactions to hallucinogenic drugs—hallucinogens occasionally produce psychotic reactions several months after use—marijuana may have been the triggering factor but this cannot be established definitely.

3. Persons with a history of psychoses.

   Many individuals who have unpleasant experiences with marijuana are ambulant schizophrenics—in some of these cases marijuana may precipitate true psychotic reactions.

Death in humans resulting directly from marijuana toxicity appears to be a rare phenomenon. Acute-toxicity determinations in animals reveal that extremely large amounts are necessary to cause death and that the median, lethal dose-to-median effective dose ratio (ie, LD50/ED50) for marijuana is many times greater than that obtained with either the barbiturates or alcohol. Children who accidently ate marijuana-containing cookies became intoxicated and presented with varying degrees of effects routinely observed in adults, eg, tachycardia, bilateral conjunctival hyperemia (congestion), ataxia and nystagmus; recovery was uneventful and occurred within 6 hr.

The continued use of marijuana may result in psychological dependence, and tolerance may develop to psychological (characteristic "high" time estimation), physiological (tachycardia) and combined (psychomotor coordination) effects of marijuana. The evidence for psychological tolerance accrues, in part, from the observation that chronic users tend to increase the amount consumed, or resort to a more potent variety in order to experience altered states of consciousness. Clinical laboratory studies provide data to support the other forms of tolerance. The mechanisms involved in tolerance to marijuana may include cellular adaptation, particularly within the CNS, and an increased biotransformation capacity.

Conversely, the phenomenon of "reverse tolerance," or sensitization to marijuana, has been reported. This may be attributable to psychological or metabolic factors, or a combination of both. Experience undoubtedly plays a role in the user's awareness and enjoyment of a marijuana-induced "high," and, with repeated conditioning, less of a stimulus is necessary to trigger the anticipated subjective effects. In addition, long-term smokers appear to be more efficient, inhaling and retaining more smoke per puff than the novice. THC and, possibly, active metabolites of this molecule are eliminated slowly from the body. Some chronic users continue to excrete THC in the urine for 20 to 30 days after terminating all marijuana smoking and/or ingestion. The frequent use of marijuana, therefore, may result in significant *in vivo* accumulation with a consequent reduction in the amount of drug needed to exceed a psychoactive threshold in the brain. Such accumulation has been reported to occur in volunteer subjects who claim having had no prior exposure to marijuana. Approximately 50% of a standardized dose of THC was present in the plasma of naive subjects 56 hr after administration. The factors possibly contributing to this prolonged retention include an enterohepatic recirculation of THC and/or active metabolites, binding to plasma proteins and sequestration in adipose tissue with delayed metabolism. In chronic marijuana smokers the biological half-life of THC was reduced appreciably (ie, $t_{1/2} = 28$ hr), but this period still is sufficiently long to result in accumulation if marijuana is used daily or more frequently.

Physical dependence may occur, since after 1 week of THC administration, a withdrawal syndrome has been observed which consisted of anorexia, nausea, insomnia, sweating, hyperthermia and tremor. The mildness of these responses probably is due to the slow elimination of THC from the body, which allows physiological and psychological systems to adjust to a drug-free state gradually.

Under experimental conditions employing male animals, and in human smokers, marijuana decreased testosterone blood levels, testicular size and weight, spermatogenesis and sexual potency. The inhibition of the release of luteinizing hormone (LH) from the pituitary gland, and the testicular responses to LH stimulation have been cited as possible mechanisms. However, THC also has weak estrogenic activity, as demonstrated by animal studies and clinical examination (including biopsy) of young males who developed gynecomastia during heavy marijuana use. THC inhibits ovulation in rats, rabbits and monkeys. The disruption of menstrual cycles has occurred in women who smoke marijuana on a regular basis.

Studies conducted with laboratory animals have shown that prolonged administration of THC may inhibit growth, impair lactation, promote thyroid hyperplasia and elevate plasma corticosteroid levels. These physiological alterations appear to reflect primarily actions of THC on the pituitary gland. High doses of THC in animals have been reported to induce hyperactivity and convulsive seizures indicative of neurotoxicity. Lacking comparable data in humans the significance of these studies must be interpreted cautiously.

Prolonged marijuana use may lead to serious pulmonary toxicity. *In vitro* tests employing lung explants demonstrated that marijuana smoke can induce premalignant and malignant cellular changes. Chronic exposure of animals to marijuana smoke led to severe bronchiolitis and squamous metaplasia of the tracheal mucosa, and fatal respiratory complications occurred in some cases. Bronchial biopsies in humans who were long-term marijuana smokers also revealed squamous metaplasia. Substantial respiratory impairment, indicated by a significant increase in resistance to airflow (suggestive of obstructive lung disease), and high carboxyhemoglobin levels also have been observed in these individuals. Both abnormalities are comparable to those associated with chronic tobacco smoking. In this regard, smoking one marijuana cigarette increases the concentration of carbon monoxide and tar in the lungs comparable to five or more tobacco cigarettes. Pulmonary toxicity should be considered a probable consequence of chronic marijuana smoking.

The suppression of cellular-mediated immune responsiveness has been demonstrated in young, chronic marijuana smokers, but this effect is transitory. The lymphocytic response observed in marijuana smokers was similar to that of patients in whom impairment of T (thymus-derived) cell immunity is known to occur. Some clinical studies have shown no significant suppression of lymphocyte function. Current clinical data indicate no increase in malignancies and infections among chronic marijuana smokers.

Personality, attitudinal and behavioral changes frequently are associated with chronic marijuana smoking. There characteristically is a reduction in motivation, the desire to be productive, creative or contributive, and the individual may experience acute feelings of insecurity. Although elements of this syndrome are typical of normal adolescent turmoil, compulsive involvement with marijuana may accelerate or project into, intensify and delay emergence from this ambivalent phase of life. Marijuana may foster similar disruptions in older persons but evidence also exists that individuals can continue to function effectively in artistic and other creative areas while indulging in frequent but moderate use of the drug.

The LaGuardia Report (Mayor's Committee on Marijuana, New York City, 1944) stated that "marijuana will not produce a psychosis *de novo* in a well-integrated stable person." Judging from the medical literature published subsequent to this report, primary marijuana psychosis is relatively rare in the US. The precipitation of serious psychological problems appears to occur primarily in persons with preexisting personality or emotional disturbances. The use of marijuana by schizophrenic patients, including those being treated with antipsychotic agents, may result in rapid and serious deterioration of their mental state necessitating rehospitalization in some cases.

Some studies have demonstrated a positive correlation between marijuana dosage and birth defects. However, other investigations have failed to provide evidence that marijuana possesses teratogenic activity. THC administered to

pregnant rats and dogs is transferred rapidly to fetal tissue and results in a higher than expected incidence of abnormal pregnancies and stillborn offspring. Malformations observed include cleft palate, accessory ribs, fused ribs and asynchronous and retarded vertebral ossification. Women who smoke marijuana while pregnant experience a longer period of labor and their newborn weigh less than normal and have altered CNS activity. THC is lipid-soluble and passes into the milk of the lactating female. Thus, marijuana specifically should be avoided by women who are breast-feeding their newborn.

Although primary attention has been directed to the adverse physiological and social effects of marijuana, there are several indications that the tetrahydrocannabinols may possess clinically useful properties. When administered to patients with advanced cancer, oral doses of THC (capsules containing 7 to 10 mg in sesame oil) elicited mild analgesic, antidepressant, tranquilizing and antiemetic effects. However, a rapid development of tolerance, sometimes by the third dose, has limited THC use in these patients. Further, at these doses, and more frequently at a higher (20 mg) dose, disturbing side effects, eg, dizziness, ataxia, blurred vision and excessive sedation, were observed. Although it often stimulates appetite, marijuana is not useful in treatment of anorexia nervosa. In fact, it probably should be contraindicated since persons with this disorder possess some underlying psychological abnormality which can be exacerbated by oral THC administration, ie, some patients receiving this therapy have developed significant dysphoria manifested as paranoia and loss of self-control. Other investigations have demonstrated significant and prolonged reduction of intraocular pressure by marijuana in glaucoma patients. The proposed anti-inflammatory and anticonvulsant activities of THC await further clinical evaluation.

Although much remains to be developed, there is beginning to emerge a reasonably clear picture of the acute pharmacological and toxicological effects of marijuana. While it will take longer to identify chronic toxic effects, the current deficiency of such observations should not, therefore, be misinterpreted.

Cigarettes

Although warnings have been published for 90 yr—"very many chronic, and often fatal, ailments are produced by the use of tobacco" (*Frank Merriwell's Book of Athletic Development*, Street & Smith, 1901)—many Americans are just recognizing the health risk from smoking cigarettes and are abandoning their use in significant numbers.

Cigarette smoking accounts for approximately ⅛ of all cancer deaths in this country and is the leading single cause of such mortality. Lung cancer and cigarette smoking have been linked convincingly by numerous clinical studies. There is a similar, though less frequent, association with pipe and cigar smoking.

Current evidence shows clearly that lung cancer deaths among women has increased substantially over the past 40 yr. This greater mortality is associated with a proportional increase in the number of women who have become cigarette smokers. Further evidence of this correlation is found in data from two states. In Washington, over a 10-yr period, the lung cancer death rate in women increased by more than 100% but the breast cancer death rate did not change significantly. In Utah, where a strong antismoking attitude prevails, the lung cancer death rate among women is less than 50% of that for breast cancer.

All smokers should be encouraged to stop since, after several years of nonsmoking, the risk of developing bronchogenic carcinoma approaches that of nonsmokers. Smokers also have a higher incidence of both periodontal disease and cancer of the oral cavity than nonsmokers. Bladder carcinoma, manifest both before and after the appearance of lung cancer, is another risk, as is cervical cancer. Switching to 'low tar, low nicotine' products may not be an improvement, since clinical studies show that smokers take more frequent and deeper puffs of these cigarettes than of regular ones in order to maintain their usual plasma levels of nicotine.

Bronchitis and respiratory tract disorders, in general, are more prevalent, not only in smokers, but among their family members as well since an exposure to cigarette smoke often is inescapable in the relatively closed atmosphere of a house or apartment.

Cardiovascular disorders occur more frequently, and the risk of death from coronary heart disease is significantly greater in smokers than in nonsmokers. In patients with hypertension, hypercholesterolemia or diabetes the risk of coronary heart disease is increased further by cigarette smoking. Peripheral vascular disease and cerebrovascular insufficiency also are encountered more often in smokers. A common link to these cardiovascular diseases appears to be the damage to blood vessel (eg, coronary artery) walls which occurs more frequently among smokers and which serves to promote formation of atherosclerotic plaques.

Myocardial infarction is a relatively rare complication in premenopausal females; however, cigarette smoking progressively increases the incidence of myocardial infarction to as much as 20-fold among women smoking 35 or more cigarettes per day. Since female hormones may be a factor in the lower rates of cardiovascular disease in women as compared to men, it is pertinent to note that menopause often occurs at an earlier age in women smokers.

Recent data also show an increase in stroke among young and middle-aged women who smoke cigarettes.

Smokers have elevated carboxyhemoglobin (COHb) levels due to inhalation of excess carbon monoxide from the combustion of tobacco. Significant carboxyhemoglobinemia reduces oxygen transport by the circulatory system. Environmental conditions result in the formation of COHb equivalent to approximately 0.5% of total hemoglobin in the nonsmoker. Smoking one pack of cigarettes per day may produce COHb in the range of 6% or more, a level which may result in interference with subtle CNS processes, eg, the judgment used in automobile driving.

Heavy smokers may show COHb levels of up to 20% of total hemoglobin, which places a substantial strain on the cardiovascular system. Such alterations in oxygen transport have led to consideration of possible restrictions on using smokers' blood for transfusions. An additional consequence of high carbon monoxide levels is secondary polycythemia, ie, tissue hypoxia due to prolonged exposure to carbon monoxide results in increased red-cell mass.

Gastrointestinal disturbances associated with smoking include epigastric discomfort, gastritis and, possibly, gastric and duodenal ulceration. An increase in gastric acid regurgitation into the esophagus apparently accounts for cigarette-induced heartburn which frequently is painful in heavy smokers.

Pyloric incompetence and subsequent reflux of duodenal juices may be a contributory factor in the gastritis and gastric ulceration since bile injures the gastric mucosa, particularly in the absence of food in the stomach. In addition, nicotine may cause areas of ischemia in the gastrointestinal tract and may reduce pancreatic buffering secretions, thus peptic ulceration may occur in the presence of even normal rates of gastric acid secretion.

Continued cigarette smoking during antiulcer therapy diminishes the probability of successful treatment.

In regard to influenza, several studies show that smokers contract this disease at a higher rate and experience a great-

er degree of incapacitation (ie, more lost work days than nonsmokers).

Considerable data show that smoking during pregnancy is associated with higher than normal rates of miscarriage, spontaneous abortion, prenatal mortality and premature birth. The newborn of women who smoke during pregnancy are more likely to be underweight, be short in stature and have a smaller head. These effects are dose-related, ie, the incidence increases in proportion to the number of cigarettes smoked per day. Weight, height and head circumference decrements persist 4 to 7 yr after birth.

Smokeless Tobacco—Switching to smokeless tobacco does *not* reduce toxicity. The use of two cans of snuff per week is equivalent to smoking two packs of cigarettes per day. The absorption of nicotine is rapid, peak plasma levels occurring within 5 minutes of application to the oral mucosa and twice as much nicotine is absorbed than from cigarettes. Leukoplakic lesions and cancers occur in the user's mouth, causing premature death in some teenagers and adults (Babe Ruth was a heavy user of smokeless tobacco and died of an oropharyngeal tumor at age 52).

Central Nervous System Stimulants

Amphetamines

The clinical indications for amphetamines include

The management of certain behavioral disturbances in children, eg, attention disorder (hyperkinetic syndrome) associated with minimal brain dysfunction.

The symptomatic control of narcolepsy.

The treatment of exogenous obesity, as short-term (ie, a few weeks) adjuncts in a regimen of weight reduction based on caloric restriction.

Benzphetamine, chlorphentermine, clortermine, diethylpropion, phendimetrazine or phentermine, alternatives to amphetamines in weight-reduction programs, also are subject to misuse and abuse. These compounds are related chemically and pharmacologically to the amphetamines, but possess a somewhat higher anorexiant-to-central stimulant ratio and peripheral sympathomimetic activity.

Misuse encompasses the episodic ingestion of amphetamines to suppress fatigue and prolong wakefulness and alertness, thus enabling the individual to continue mental or physical activity beyond his or her usual limit of endurance. Teachers frequently are witness to the futility of hyperamphetaminization—in the form of the tense, distraught student whose effective functioning is precluded by disorientation and mental short-circuiting or in the form of the exhausted and depressed student whose chemical props have collapsed.

Despite the hazards involved, long-distance truck drivers similarly use amphetamines to dispel monotony and boredom. Although the practice is overtly pernicious, the administration of amphetamines prior to engaging in athletic activity (eg, swimming, running, weight throwing) may improve performance to a degree that could be decisive in competition.

There remains a significant "gray area" of misuse—the prescribing of amphetamines and amphetamine-like drugs for unjustifiable reasons or, at best, in cases where the therapeutic rationale is questionable. To the busy medical practitioner, CNS stimulant and depressant drugs may provide an expedient, if less than ideal, means of helping his patients cope with the pressures and frustrations of everyday life. In the treatment of obesity these drugs provide a questionably effective and often self-deceptive approach to a complex biomedical problem.

Clearly, those engaged in prescribing and dispensing drugs must exercise skilled judgment in eliminating as candidates for amphetamine therapy those patients so emotionally predisposed as to explore the secondary values of their anorexiants, ie, the mental lift, elan and psychic crutch upon which they increasingly may depend to cope with crises, real or imaginary.

Amphetamine abuse relates primarily to the nonsupervised ingestion or injection of large doses of amphetamine or its many chemical derivatives to experience the drug-induced psychic excitation, euphoria or "high," and the physical maelstrom of restless energy. Methamphetamine (methedrine, "speed") is a favored congener among habitual amphetamine users who generally inject the drug into a vein. This provides an almost instantaneous onset of the euphoric effect (the "flash" or "rush") which is ineffable and ecstatic.

A marked degree of tolerance to the amphetamines can be acquired as, eventually, several grams of drug per day may be consumed. There have been reports of the use of more than 10 g of methamphetamine intravenously over a 24-hr period. Tolerance does not develop uniformly to all the CNS effects. The compulsive user may evidence increased nervousness, anxiety and persistent insomnia as the dose is increased.

In a typical pattern of abuse immense doses of amphetamines are injected every few hours around the clock. These "runs," during which the individual remains awake continuously, generally last 3 to 6 days but may be prolonged to weeks if the user is able to sleep even as little as 1 hr a day. The appetite for food is suppressed and there is a feeling of unbridled energy and a compulsion for constant activity. Intravenous injection of enormous doses of amphetamines elicits a "chemically generated trauma," which appears linked inseparably to the acquired psychological dependence. The intense psychotoxic syndrome ultimately forces an interruption of drug use and the individual lapses into a protracted period of deep sleep (the "crash").

Although it generally is considered that the amphetamines do not induce a physical dependence, abrupt withdrawal is characterized by lethargy and profound depression, both psychic and physical, which reinforces the drive to resume their use.

Massive abuse of amphetamines frequently leads to considerable mental and physical deterioration. Intravenous injection of large doses is extremely disabling, socially and psychologically, and has resulted in psychiatric complications ranging from subtle personality changes to paranoid psychoses. Harm to the individual and to society often arises during psychotoxic episodes. In contrast to the decreased psychological drives of the opiate user, the compulsive user of CNS stimulants has exaggerated drives. Analyzing the personality factors which underlie the preferential abuse of CNS stimulants versus narcotics, it has been postulated that the amphetamine abuser uses the stimulant as one of a variety of compensatory maneuvers to maintain a posture of active confrontation with the environment. In contrast, the heroin abuser reduces anxiety via repression and withdrawal.

The hyperactivity, the compulsivity, the feeling of great muscular strength, the paranoid delusions and the auditory and visual hallucinations may combine to make the amphetamine or cocaine user capable of committing serious antisocial acts. Chronic users of stimulant drugs also are accident-prone, since they are unaware of their fatigue until it overcomes them at an inopportune time.

As in any situation in which hypodermic equipment is shared without proper sterilization, there exists a risk of bloodborne infection, notably viral hepatitis and AIDS. Among amphetamine abusers, evidence has been noted of hepatic damage so common as to suggest the possibility of a direct toxic effect on the liver.

Parenteral administration of large doses of sympathomimetic amines may result in morbidity or mortality due to

intracranial hemorrhage or cardiac arrhythmias secondary to severe hypertension. Necrotizing angiitis was observed in Rhesus monkeys given repeated injections of methamphetamine for a 2-week period, and clinical descriptions of cerebral vasculitis and hemorrhage following the injection of this sympathomimetic amine have been reported. Intravenous injection of amphetamines may result in a syndrome characterized by fever, leukemoid reaction, disseminated intravascular coagulation and rhabdomyolysis. These factors may be responsible for the development of acute renal failure in certain amphetamine abusers.

MDMA (3,4-methylenedioxymethamphetamine), also known as "Ecstasy," is a "designer" drug which, according to its users, increases their awareness and the ability to communicate. In regard to toxicity, a recent study demonstrated that MDMA selectively damaged central (brain) nerve fibers in monkeys after only 4 consecutive days of administration. Since monkeys also are sensitive to MPTP, a known neurotoxin in drug addicts, this preclinical investigation suggests that humans may be at risk following MDMA use.

Cocaine

Cocaine, as extracted by chewing leaves of the coca plant (*Erythroxylon coca*), has dispelled hunger, provided a sense of well-being and enhanced the physical endurance of Andean Indians since before the Conquistadors. Even today, in the Andean regions of South America, chewing coca leaves is regarded as no more deviant a practice than smoking tobacco leaves by persons in other parts of the world.

The subjective effects, toxicity and present-day patterns of cocaine abuse are remarkably similar to those of amphetamine. Until recently, cocaine was very expensive when purchased from illicit sources. However, larger amounts are now being smuggled successfully into the US, leading to reductions of the "street" price. This lower cost, in the presence of a more plentiful supply, has resulted in a greater number of citizens becoming cocaine addicts. When it is unavailable, abusers often resort to amphetamine. Extemporaneous mixtures of cocaine and amphetamine or heroin are common in the contemporary drug scene.

Regardless of the route of administration of cocaine (oral, nasal insufflation, intravenous), there is good correlation between the appearance of certain physical effects (tachycardia, elevated blood pressure) and psychological alterations ("high," pleasantness, anorexia). Free-base cocaine available as "crack," is absorbed rapidly after smoking; peak plasma levels occur within minutes.

Prolonged use may be associated with weight loss, insomnia, anxiety, paranoia, sensations of insects crawling under the skin ("cocaine bugs") and hallucinations (primarily visual—flashes of light or "snow lights"; may also be tactile, olfactory and auditory). Ulceration and perforation of the nasal septum also may occur. In one reported case of chronic cocaine sniffing, the patient presented with a continuous nasal discharge that was not mucus. Instead, it was shown to be cerebrospinal fluid leaking from the CNS area due to extensive cocaine-induced local tissue and nerve (olfactory) damage.

Large doses of cocaine may result in cardiac dysrhythmias, tremors, convulsions and delirium. Deaths have been reported following every route of cocaine administration, including nasal insufflation. Unusual fatalities have occurred in drug dealers who, to avoid detection, swallowed prophylactics filled with cocaine; when several condoms ruptured in the gastrointestinal tract, lethal concentrations of cocaine were absorbed.

Tolerance to cocaine develops very rapidly (tachyphylaxis), particularly when used daily. Although a "line" of cocaine has about 25 mg, some addicts have used 8 to 9 *grams* per day. Treatment consists of abrupt and complete cessation (as opposed to gradual—approximately 7 days—reduction with most CNS depressants).

A withdrawal syndrome, which includes increased appetite, fatigue (abuser may sleep for 24 straight hr) and depression (with increased suicidal tendency) usually occurs in cases of chronic administration. The craving for cocaine during withdrawal is very intense during the first 7 days and appears to be linked to hypersensitive dopamine receptors (compensatory biological adaptation to cocaine-induced dopamine depletion). Bromocriptine (Parlodel), a dopamine receptor agonist, has been employed successfully in treating this aspect of cocaine withdrawal.

Physical dependence, therefore, does occur with chronic cocaine abuse. However, its presence is unnecessary when classifying someone as an addict since addiction is characterized as "a behavioral pattern of compulsive drug abuse" associated with "overwhelming involvement with the use of a drug, the securing of its supply and a high tendency to relapse after withdrawal.[2]" In this frame of reference, the chronic user of cocaine or amphetamine is an *addict*.

Psychotomimetics

Psychotomimetics constitute a structurally diverse group of naturally occurring and synthetic molecules. Interest in these compounds resides more in their misuse than in their legitimate medical use. They are of value as research tools in experimental psychiatry and in the exploration of central neurochemical mechanisms, but their therapeutic application remains limited and highly controversial.

At high dosage levels many drugs may disorganize mental function with resulting confusion, delirium, hallucinations and, frequently, memory loss or amnesia. Such drugs include atropine, scopolamine (and related centrally acting anticholinergics), quinine, quinidine, digitalis glycosides, mecamylamine, adrenocorticosteroids, nalorphine, disulfiram, bromides and certain heavy metals. The *toxic psychoses* produced by these drugs are due primarily to generalized metabolic disruption of both neural and extraneural systems rather than to discrete neurophysiological perturbations.

Certain chemicals, however, are uniquely capable of inducing dramatic changes in psychic processes (ie, perception, thought, feeling, mood and behavior) in doses which do not produce generalized metabolic disruption and which do not cause marked disturbances in sensorimotor or autonomic functioning. These compounds generally are classified as *psychotomimetics*, although the extent to which they mimic spontaneously occurring psychotic states is inconsistent and incomplete. Other imaginative designations for such substances include *psychosomimetics*, *psychotogenics*, *psychodysleptics*, *psychedelics*, *hallucinogenics*, *mysticomimetics* and *phantasticants*.

On a structural basis, psychotogenic chemicals may be classified into three major groups:

Substituted *indole alkylamines*, eg, dimethyltryptamine, psilocybin or lysergic acid diethylamide.

Substituted *phenyl alkylamines*, eg, mescaline or dimethoxymethylamphetamine.

A structurally *heterogeneous* group, including the glycolate ester, ditran [a mixture of *N*-ethyl-3-piperidyl-(30%) and *N*-ethyl-2-pyrrolidyl-methylcyclopentylphenyl glycolate (70%)] and the piperidine derivative, phencyclidine.

With the exception of lysergic acid diethylamide, the chemical nature and pharmacological properties of the various psychotomimetics will be considered only briefly. The interested reader is referred to several comprehensive reviews on this extensive and complex category of psychoactive agents (refer to the *Bibliography*).

Dimethyltryptamine

Hallucinogenic activity is characteristic of a large series of *N*-alkylated tryptamines. Structurally, the simplest of these is *N,N*-dimethyltryptamine (DMT). This compound occurs naturally in the seeds of *Piptadenia peregrina*. A powder prepared from these seeds, and referred to as *cohaba snuff*, is used by Haitian natives to induce mystical states of consciousness. DMT is not effective when taken orally. Perceptual and mood changes result when the compound is inhaled (snuffed), smoked or introduced parenterally. Its effects are rapid in onset and limited in duration (a few hours), irrespective of the route of administration. Synthetic higher homologs of DMT, ie, diethyltryptamine (DET) and dipropyltryptamine (DPT), produce qualitatively similar psychological effects which are, however, considerably longer-lasting.

Psilocybin and Psilocin

Psilocybin, the phosphate ester of 4-hydroxy-DMT occurs to the extent of about 0.3% in the Mexican mushroom, *Psilocybe mexicana*. Dephosphorylation *in vivo*, by alkaline phosphatase, converts psilocybin to psilocin (4-hydroxy-DMT). The loss of the phosphoric acid radical reduces the polarity of the molecule, enabling more-efficient penetration of the blood-brain barrier, which may account for the relatively greater hallucinogenic potency of psilocin as compared to psilocybin. Although psilocin is less potent than LSD (ie, approximately $^1/_{100}$ as active on a milligram basis) and produces a less-persistent psychedelic state, when equivalent doses are administered blind it generally is impossible for subjects acquainted with the LSD phenomenon to differentiate between the two drugs.

Mescaline

One of the first phenyl alkylamine hallucinogens to be identified was mescaline (3,4,5-trimethoxyphenethylamine), isolated originally from "mescal buttons," the flowering heads of the peyote cactus, *Lophophora williamsii*. This plant material long has been used by the Mescalero Apaches of the Southwest American plains in their quasireligious ceremonies of peyotism. Mescaline is not a particularly potent psychotomimetic. The equivalent oral dose of mescaline (usually 5 mg/kg in humans) is approximately 4000 times larger than that of LSD. Following oral administration, mescaline produces a characteristic syndrome of sympathomimetic effects, anxiety, hyperreflexia, static tremors and psychic perturbations including vivid hallucinations, usually of a visual nature. In man, mescaline has a biological half-life of about 6 hr. It is excreted in the urine principally in the form of the unaltered drug and the inert metabolite 3,4,5-trimethoxyphenylacetic acid.

The addition of an alpha-methyl substituent to mescaline produces 3,4,5-trimethoxyamphetamine (TMA), a psychotogen approximately twice as potent as mescaline. Its enhanced potency is due presumably to a decreased susceptibility to oxidative deamination provided by alkylation of the alpha-carbon.

The TMA analogue, 2,5-dimethoxy-4-methylamphetamine (DOM), is a potent psychedelic agent employed extensively by certain drug abusers and designated by them as STP (an acronym derived ostensibly from the terms "serenity, tranquility, peace"). In doses of 5 mg or more, it produces intense and relatively long-lasting emotional changes and perceptual distortions. Cases have been reported of individuals actively hallucinating for several days following a single oral dose.

The consideration of the pharmacology and structure-activity relationships of the numerous synthetic dimethoxyamphetamines, trimethoxyamphetamines and methoxymethylenedioxyamphetamines is beyond the scope of this presentation; this area has been reviewed extensively by Shulgin *et al*[3] and Snyder and Richelson.[4]

Lysergic Acid Diethylamide

The dextrorotatory isomer of lysergic acid diethylamide (LSD), synthesized by Hofmann in 1938, remains the most potent psychotogenic agent either of natural or synthetic origin discovered to date. Although as little as 25 μg of LSD may produce subjective effects, intense depersonalization usually requires doses in the range of 100 to 250 μg. Structurally, LSD is related to the ergot alkaloids, notably ergonovine. This structural resemblance may account for certain pharmacological and toxicological similarities among LSD and the lysergic acid alkaloids of ergot.

Metabolism—Following oral administration, LSD is absorbed rapidly and widely, but not distributed uniformly throughout the body. It is bound strongly to plasma proteins; highest concentrations are found in the liver, kidneys and lungs. Considerably less than 1% of an orally administered dose penetrates into the CNS. Autoradiographic analyses of brain samples obtained from animals injected with ^{14}C-labeled LSD revealed relatively high concentrations in the auditory and visual reflex areas. While the distribution of LSD within the brain would appear to suggest the functional involvement of specific neural areas in the psychotogenic phenomenon, there is an imperfect correlation between drug localization and sites of drug action.

In humans the biological half-life of LSD is approximately 3.5 hours; this corresponds roughly with the duration of the peak psychosensory effects which then subside gradually over an 8- to 12-hr period.

Pharmacological Effects—LSD possesses considerable CNS-stimulant activity. It produces an EEG pattern characteristic of central activation, alertness or arousal and causes insomnia in laboratory animals and humans. LSD counteracts the central depressant effect of barbiturates and is antagonized by such suppressants as chlorpromazine.

LSD produces a sequential, though somewhat overlapping, pattern of physiological and behavioral changes, the intensity and duration of which largely are dose-dependent. Pupillary dilation, tachycardia, tremulousness, hyperthermia and elevated blood glucose and free-fatty-acid levels, indicative of adrenergic activation, frequently are manifest during the early phases of the LSD response. These physiological alterations may be attributed both to primary LSD effects and to nonspecific stress-anxiety reactions.

Controlled studies of individuals under the influence of LSD uniformly reveal a generalized impairment of objective indices of adaptive behavior and psychomotor performance, especially those processes and procedures that require critical judgment and coordination. It is likely that intellectual and motor decrements are due to attenuation of attention and motivation as well as to sensory-cognitive disturbances.

Perceptual alterations constitute the most dramatic effects of LSD; their kaleidoscopic patterns defy a brief description. Illusions and pseudohallucinations, mostly of a visual or tactile nature, are experienced commonly, whereas true hallucinations are relatively infrequent. Synesthesia, the crossover from one sensory modality to another, is an often-encountered LSD phenomenon. Colors may be "heard" and music may become "palpable." Moods and emotions may range from euphoria, elation and ecstasy to dysphoria, depression and despair. The psychological state produced by LSD cannot be generalized with precision. As with other psychotropic drugs, the response depends on many variables, including the dose administered, the personality and expectations of the individual as well as environmental influences.

Mechanisms of Action—The neurophysiological corre-

lates of LSD-induced alterations in behavior are understood incompletely. However, recent data indicate that LSD and other hallucinogens act at postsynaptic serotonin receptor sites ($5HT_2$ subtype). The effect of LSD upon raphe neurons resembles that of an excess of serotonin at postsynaptic receptor sites.

Experimental and Therapeutic Uses—LSD has been employed extensively to induce experimental psychoses for the primary purpose of studying aberrant mental states under controlled conditions. Despite prodigious efforts, the LSD model has not yielded pertinent clues to the biochemical etiology of schizophrenia.[4]

Several investigators have proposed LSD as an adjunct to conventional psychotherapy and as an aid in treatment of chronic alcoholism. LSD also has been reported to provide long-lasting "euphor-analgesia" in patients with terminal cancer. The feasibility and effectiveness of LSD for these purposes remain unestablished and controversial. LSD has no approved therapeutic uses and currently is an investigational drug subject to rigid state and federal regulations.

Dependence Liability—Marked psychological dependence on LSD is observed rarely as usage tends to be occasional or sporadic rather than frequent or compulsive. A high degree of tolerance to the physiological and behavioral effects of LSD develops after three or four doses taken within a relatively short period of time. This acquired resistance disappears rapidly if drug intake is terminated. There is considerable cross-tolerance among LSD, mescaline and psilocybin, but this phenomenon has not been demonstrated between LSD and either amphetamine or Δ^9-THC. As physical dependence on LSD does not develop, there is no characteristic abstinence syndrome upon abrupt discontinuation.

Toxicity—Despite its extreme psychotogenic potency the acute toxicity of LSD is remarkably low. The medical literature records no verified case of death in man attributable to the direct toxic effects of the drug, although fatal accidents and suicides have occurred during states of LSD intoxication. Homicides committed by persons apparently under the influence of LSD have been reported relatively infrequently. Most of the individuals involved evidonced premorbid psychopathological tendencies and thus the role of LSD in violent and assaultive behavior is equivocal.

LSD-induced feelings of depersonalization and affective, perceptual and cognitive distortions may, on occasion, result in disorientation, confusion and acute panic reactions characterized by anxiety, fear and a sense of helplessness and loss of control. "Bad trips" generally follow the ingestion of high doses of LSD by nontolerant persons. They also are likely to occur in inexperienced users, those with ambivalent attitudes toward the drug experience or in disturbing or threatening surroundings. Reequilibration usually takes place within 24 to 48 hr.

Recurrences of perceptual distortions may be experienced in the postdrug state by a relatively high percentage of LSD users. These "flashbacks," which vary in length from a few seconds to several minutes, may occur up to 5 yr after the drug was last taken. Flashbacks may be spontaneous but often are triggered by periods of emotional stress or anxiety or by other psychotropic drugs, such as marijuana. The mechanism of recurrent hallucinosis is unknown but may reflect a persistent disruption of psychological defense mechanisms with a periodic emergence of repressed fears or conflicts.

Chronic disruptive states associated with anxiety, depression, somatic disturbances and difficulty in functioning, which are relatively resistant to psychotherapy, commonly follow LSD use. Protracted schizophreniform psychotic states with paranoid behavior represent infrequently occurring but tragic psychological consequences of LSD. Most, but possibly not all, such cases involve unstable individuals with prepsychotic or premorbid personality traits. An unfavorable prognosis is indicated by motor retardation, withdrawal, blunt affect, anergy and suicidal ideation during the initial hospitalization period. Treatment varies, but lithium has been proven effective for the alleviation of LSD-induced psychosis.

There are several reports of inflammatory fibrosis occurring in individuals who have consumed LSD. This complication has been recorded previously with other lysergic acid derivatives, notably methysergide. Arteriospasm resulting in obstruction of the internal carotid artery, and the development of peripheral gangrene necessitating partial amputation of the extremities, constitute isolated case reports indicating that LSD shares the vasoconstrictor activity of other ergot alkaloids.

In 1967 investigators first reported chromosome damage in human leukocytes cultured *in vitro* with LSD. Although the clinical significance of this finding was exaggerated grossly in the public news media, the widespread publicity contributed to a significant downturn in the abuse of LSD at that time. The possibility of affecting generations yet unborn apparently struck a chord of moral responsibility in many who were convinced of their personal ability to maintain psychic control but who were unwilling to "pollute the genetic stream."

Genetic studies conducted with LSD have been reviewed critically by Dishotsky *et al.*[5] Although the relationships between LSD and chromosomal damage, leukemogenicity and teratogenicity remain unresolved, certain tentative conclusions appear warranted.

Data supporting a positive relationship between LSD and chromosomal aberrations have been obtained primarily with individuals reported to have taken LSD obtained in the black market. In most instances, the amount of LSD consumed cannot be ascertained or only can be approximated. The reputed LSD samples may contain other drugs or contaminants, either added or incompletely separated during the process of illicit synthesis. The population under study frequently extemporize with barbiturates, amphetamines, opiates, cocaine, marijuana and other psychotogens, in addition to LSD.

Chemically pure LSD administered under controlled conditions has, in several studies, failed to produce detectable damage to chromosomes or has produced transient chromosomal aberrations in peripheral leukocytes, but these defects were no longer evident several months after LSD administration. Transient chromosomal breaks in white blood cells occur spontaneously. They can be increased by certain antibiotics and antineoplastic agents and even by commonly employed drugs such as aspirin and caffeine. Viral infections are associated with an increased rate of chromosomal disruption. Hepatitis, gastrointestinal and upper respiratory viral infections are common among chronic drug abusers. Thus, it appears that chromosomal damage, when found, is related to a history of drug abuse in general and not to LSD specifically.

The pathological significance of chromosomal aberrations in continuously replenished peripheral leukocytes is equivocal. Testicular and bone-marrow biopsies in rhesus monkeys given repeated oral doses of LSD have not revealed significant chromosomal alterations in gametogenic and hemopoietic tissues.

Two cases of acute leukemia developing subsequent to the use of LSD are recorded. Although a causal relationship has not been established it may be premature to dismiss the association as merely coincidental.

Some studies suggest a higher incidence of spontaneous abortion among pregnant women who reportedly took LSD prior to or after conception, and a greater number of congen-

ital anomalies among live infants born to mothers exposed to this drug. However, several complicating factors preclude a definitive correlation of increased reproductive risk with LSD ingestion. Among these are the indeterminate nature of purported LSD samples obtained "on the street," a common history of multiple usage of illicit drugs, a high incidence of infectious diseases (especially viral illnesses) and marginal maternal nutrition. Although the effect of LSD on human pregnancy and fetal malformations remains uncertain, discretion dictates the avoidance of this drug by women of childbearing age.

Phencyclidine

Phencyclidine (PCP, "angel dust"), chemically and pharmacologically similar to ketamine (Ketalar) used to induce "dissociative anesthesia," is probably the most dangerous substance abused in the US. There is no consensus as to the precise pharmacological classification of PCP. The compound may, depending on the dose and other circumstances of use, exhibit stimulant, depressant, analgesic and hallucinogenic properties. In "street" form, PCP often is adulterated and frequently misrepresented as THC, mescaline, LSD, amphetamine, cocaine or many other psychoactive agents.

Although occasionally ingested orally or injected intravenously, PCP most commonly is smoked (after placing it on marijuana or dried parsley leaves in a "joint") or "snorted" (nasal insufflation). By smoking, the experienced user can limit the dose of PCP (self-titration) to a level with which he is comfortable and less likely to overdose than when the drug is taken orally.

While PCP ingestion can produce euphoria, adverse reactions more commonly are observed, particularly in naive users. An excellent classification of PCP effects has been developed by Rappolt *et al*[6] based upon their treatment of more than 250 cases. Tachycardia and elevated blood pressure are consistent findings and appear, in varying degrees, within each of the following categories:

Stage I: 2 to 5 mg PCP (serum concentration, 25 to 90 ng/mL)

Subjects are disoriented, combative and violent. They also experience ataxia, alterations in perception of visual, auditory and tactile sensations, excessive sweating and salivation and analgesia (they may injure themselves unknowingly due to this analgetic property).

Deaths occurr when subjects lose control of motor function yet attempt activities which require significant physical skill, eg, some try to swim but subsequently drown. Other fatalities happen after abusers engage in violent fights or fall asleep in the middle of a street and are crushed by a motor vehicle.

Stage II: 5 to 25 mg PCP (serum concentration, 90 to 300 ng/mL)

The patient presents with coma and does not respond to verbal communication; reactions to painful stimuli will occur, however. Muscle spasms and severe hyperthermia also may be present.

Stage III: Above 25 mg PCP (serum concentrations, above 300 ng/mL)

Deep coma is observed with patients showing no response to extremely painful stimuli. Seizures also are likely and may develop into status epileptics.

Although the data are more difficult to interpret, it appears that a number of deaths solely and directly are related to excessive blood levels of PCP. Cerebral hypoxia due to severe spasm of cerebral blood vessels may be a mechanism of lethality.

Delayed psychological reactions (delirium, psychosis and/or agitation) occurring approximately 1 week after consumption of high doses of PCP have been observed. This may be due to the high, lipid-solubility of the drug resulting in an accumulation in, and slow release from, adipose tissue; the $t_{1/2}$ is approximately 18 hr. On occasion, patients hospitalized for a psychiatric examination have their blood analyzed for PCP levels. In some of these cases, a result showing an absence of PCP may be incorrect. The methods of analysis using high-performance liquid chromatography (HPLC), gas chromatography with flame ionic detection (GC-FID) or radioimmunoassay (RIA) are accurate only down to levels of 100 to 200 ng/mL. However, as presented above, serum PCP concentrations between 25 and 90 ng/mL are sufficient to induce aberrant behavior. A recent study employing a more sensitive assay procedure, a glass capillary-gas chromatography thermionic specific (nitrogen) detector (GC^2-N) capable of measuring levels as low as 5 pg/mL, reported that of 135 patients admitted for psychiatric evaluation, 78 had PCP levels between 1 and 50 ng/mL. This is a significant observation since it can assist physicians in determining the correct treatment.

A two- to four-fold tolerance develops if PCP is administered chronically to laboratory animals. However, experiments performed to date do not suggest that PCP produces physical dependence comparable to that which develops to the opiates or other CNS depressants.

In normal volunteers, PCP induces a schizophrenic-like state. Thus, as is the situation with marijuana, individuals with psychoses (diagnosod or undiagnosed) particularly are vulnerable to PCP. Schizophrenics experience a deterioration of their condition, possibly culminating in stuporous or excitatory catatonia or paranoia accompanied by auditory hallucinations.

Rhabdomyolysis (skeletal muscle degeneration), myoglobinuria and renal failure have developed after acute, large doses of PCP, whereas chronic use is associated with both psychological and physical dependence, and alterations in memory, speech and vision. These latter changes are suggestive of organic brain damage.

Treatment of Acute Drug Overdosage

A major problem in treating incoherent drug-overdosed patients, ranging from comatose to delirious, is the absence of definitive data regarding the substance(s) responsible for the intoxication. Upon admission to an emergency center it is imperative that staff members consult persons on the scene or the patient's friends in an attempt to obtain as much information as possible about the drug(s), amounts and modes of administration, circumstances leading to the overdosage and pertinent aspects of the patient's medical history, eg, does the patient have diabetes or epilepsy? Due, however, to extensive adulteration of "street" drugs, the information obtained on drug identity and quantity must be evaluated with caution. Symptomatic treatment is advisable until a definitive diagnosis can be established. The following is a limited presentation of options available for treating adverse reactions to psychoactive substances.

Volatile Hydrocarbons—The treatment of acute intoxication with volatile hydrocarbons is similar to that employed for barbiturate overdosage. If the vapors are inhaled, oxygen (or a 95% O_2 and 5% CO_2 gas mixture) may be administered. When volatile hydrocarbons are swallowed, gastric lavage rather than an emetic should be used. The injection of epinephrine or other sympathomimetic amines should be avoided due to the possibility of myocardial sensitization and precipitation of cardiac arrhythmias.

Opioids—Naloxone remains the drug-of-choice in countering narcotic analgetic overdosage. This narcotic antagonist, which possesses little or no agonistic activity, may be administered to the unconscious patient in the absence of a definitive diagnosis of narcotic overdosage. Naloxone will not produce additional CNS-depressant effects in the event that acute poisoning is due to barbiturates or other nonnarcotic depressants.

Psychotomimetics—In cases of adverse psychological reactions to hallucinogens ("bad trips"), patients should be maintained in a supportive and nonthreatening environment. Verbal contact should be established for reality defining and reassurance ("talk-down") that the episode eventually will terminate. If pharmacological intervention appears indicated, the use of diazepam (or a related benzodiazepine derivative) avoids the hazards which may be encountered with a phenothiazine in an unsuspected case of anticholinergic drug intoxication or in an individual with a history of convulsive disorders. When known anticholinergic agents are taken in excessive quantities, physostigmine, which antagonizes both central and peripheral atropine-like effects, is the drug-of-choice.

Phencyclidine—The treatment of PCP overdosage differs from that associated with hallucinogens as intoxicated patients should not be engaged in an extended "talk-down" process. Isolation, with periodic observation, is beneficial as in relieving the symptoms of acute schizophrenic reactions. Diazepam may control severe agitation. Acidification of the urine with ascorbic acid or cranberry juice (avoid ammonium chloride or orange juice) accelerates the excretion of PCP and may reduce the incidence of delayed reactions.

Cocaine—Adverse reactions to cocaine are usually of short duration and may terminate before treatment is initiated. Propranolol may be employed to attenuate the cardiovascular disturbances in cases of moderate cocaine overdosage. Diazepam may suppress the CNS excitation, although the possibility of adding to subsequent cocaine-induced respiratory depression must be considered.

Amphetamines—Disturbances of the sympathetic nervous system observed in amphetamine toxicity should be treated if they threaten the patient. Acidification of the urine (avoid ammonium chloride or orange juice) can shorten the duration of attendant psychoses significantly. In the presence of acute renal failure accompanying shock and rhabdomyolysis associated with amphetamine intoxication, substantial fluid replacement is indicated.

* * * *

Pharmacists can participate in the early management of acute drug poisoning by advising the use of ipecac *syrup* (not the fluidextract) in appropriate situations. If the subject has ingested a potentially harmful quantity of drugs and is conscious, *syrup of ipecac* may be employed in the following oral doses: patient under 1 yr—10 mL; 1 to 12 yr—15 mL; over 12 yr—30 mL. Subsequently, 250 to 500 mL of liquid should be given. Vomiting within 30 min occurs in approximately 90% of patients receiving this regimen. If emesis does not ensue within 30 min, the recommended dose, with additional fluids, may be repeated. Syrup of ipecac is less useful if more than 60 min have elapsed since consumption of the drug overdose. If the patient does not vomit after two doses of the ipecac, the dosage should be recovered by gastric lavage.

References

1. Weil AT: *N Engl J Med* 282: 997, 1970.
2. Jaffe JH. In Gilman AG, Goodman LS *et al*, eds: *The Pharmacological Basis of Therapeutics*, 7th ed, Macmillan, New York, 532, 1985.
3. Shulgin AT *et al*: *Nature* 221: 537, 1969.
4. Snyder SH, Richelson E. In Efron DH, ed: *Psychotomimetic Drugs*, Raven, New York, 43, 1970.
5. Dishotsky NI *et al*: *Science* 172: 431, 1971.
6. Rappolt RT *et al*: *Clin Toxicol* 16: 509, 1980.

Bibliography

Opioids

Cami J *et al*: *Clin Pharmacol Ther* 38: 336, 1985.
Christie DJ *et al*: *Arch Int Med* 143: 1174, 1983.
D'Agostino and Arnett EN: *JAMA* 241: 277, 1979.
Nathenson G *et al*: *J Pediatr* 81: 899, 1972.
Pearson MA *et al*: *Clin Toxicol* 15: 267, 1979.
Singer TP *et al*: *Toxicology* 49: 17, 1988.
Soin JS *et al*: *JAMA* 224: 1717, 1973.

Barbiturates

Kay DC *et al*: *Clin Pharmacol Ther* 13: 221, 1972.
Smith DE, Wesson DR: *JAMA* 213: 294, 1970.
Solomon F *et al*: *N Engl J Med* 300: 803, 1979.

Nonbarbiturate Depressants

Faulkner TP *et al*: *Clin Toxicol* 15: 23, 1979.
Gerald MC, Schwirian PM: *Arch Gen Psychiat* 28: 627, 1973.
Ionescu-Pioggia, M *et al*: *Int Clin Psychopharmacol* 3: 97, 1988.
Schnoll SH, Fishkin R: *J Psychedel Drugs* 5: 79, 1972.

Alcohol

Agarwal BN *et al*: *NY J Med* 73: 1331, 1973.
Becker JT *et al*: *Alcohol: Clin Exper Res* 7: 213, 1983.
Gill JS *et al*: *N Engl J Med* 315: 1041, 1986.
Hammond KB *et al*: *JAMA* 226: 63, 1973.
Harris RE, Wynder EL: *JAMA* 259: 2867, 1988.
Jones KL *et al*: *Lancet* 1: 1267, 1973.
Jones RJ *et al*: *JAMA* 249: 2517, 1983.
Myers RD: *Ann Rev Pharmacol Toxicol* 18: 125, 1978.
Slavney PR, Gran G: *J Clin Psychiatry* 39: 782, 1978.

Volatile Hydrocarbons

Hayden JW *et al*: *Clin Toxicol* 11: 549, 1977.
Haverkos HW, Dougherty J: *Am J Med* 84: 497, 1988.
Sigell LT *et al*: *Am J Psychiatry* 135: 1216, 1978.

Aerosols

Harris WS: *Arch Intern Med* 131: 162, 1973.
Kilen SM, Harris WS: *J Pharmacol Exp Ther* 183: 245, 1973.
Sharp CW, Brehm ML, eds: *NIDA Research Monograph* 15: October, 1977.
Steadman C *et al*: *Med J Aust* 141: 115, 1984.

Marijuana

Dalterio S *et al*: *Pharmacol Biochem Behav* 8: 673, 1978.
Gross H *et al*: *J Clin Psychopharm* 3: 165, 1983.
Hollister LE: *Pharmacol Rev* 38: 1, 1986.
Lemberger L, Rubin A: *Drug Metab Rev* 8: 59, 1978.
Reeve VC *et al*: *Drug Alcohol Depend* 11: 167, 1983.
Tashkin DD *et al*: *Ann Intern Med* 89: 539, 1978.
Treffert DA: *Am J Psychiatry* 135: 1213, 1978.
Weinberg D *et al*: *Pediatrics* 71: 848, 1983.
Wu T-C *et al*: *N Engl J Med* 318: 347, 1988.

Cigarettes

Colditz, GA *et al*: *N Engl J Med* 318: 937, 1988.
Edwards TA: *Br Med J* 1: 637, 1977.
Fielding JE: *N Engl J Med* 298: 337, 1978.
Grimes DS, Goddard J: *Br Med J* 2: 460, 1978.
Jick H *et al*: *Lancet* 1: 1354, 1977.
Kline J *et al*: *N Engl J Med* 297: 793, 1977.
Read NW, Grech P: *Br Med J* 3: 313, 1973.
Slone O *et al*: *N Engl J Med* 298: 1273, 1978.
Stolley PD: *Ibid* 309: 428, 1983.
Trevathan E *et al*: *JAMA* 250: 499, 1983.

Smokeless Tobacco

Benowitz NL *et al*: *Clin Pharmacol Ther* 44: 23, 1988.

Central Nervous System Stimulants

Cohen S: *JAMA* 231: 74, 1975.
Dackis CA, Gold MS: *Lancet* 1: 1151, 1985.
Finkle B, McCloskey KL: *J Forensic Sci* 23: 173, 1978.
Goldstein FJ: *Med Times* 116: 123, 1987.
Resnick RB, Kestenbaum RS: *Science* 195: 696, 1977.
Ricaurte GA *et al*: *JAMA* 260: 51, 1988.
Sawicka EH and Trosser A: *Br Med J* 286: 1476, 1983.
Siegl RK: *Am J Psychiatry* 135: 309, 1978.
Suarez CA *et al*: *JAMA* 238: 1391, 1977.

Psychotomimetics

Brawley P, Duffield JC: *Pharmacol Rev* 24: 31, 1972.
Efron DH, ed: *Psychotomimetic Drugs*, Raven, New York, 1970.
Jacobs BL: *Am Sci* 75: 386, 1987.

Lake CR *et al*: *Am J Psychiatry 138:* 1508, 1981.
Lipton MA *et al*, eds: *Psychopharmacology: A Generation of Progress*, Raven, New York, 1978.
Smythies JR. In *Neurosciences Research Symposium Summaries*, MIT Press, Cambridge, MA, 1971.
Uadin E, Forrest IS, eds: *Psychotherapeutic Drugs*, Dekker, New York, 1976–77.

Phencyclidine

Allen RM, Young SJ: *Am J Psychiatry 135:* 1081, 1978.
Aniline D *et al*: *Biol Psychiatry 15:* 813, 1980.
Burns RS *et al*: *West J Med 125:* 345, 1975.
Burns RS, Lerner SE: *Clin Toxicol 12:* 463, 1978.
Cohen S: *JAMA 238:* 515, 1977.
Cook CE *et al*: *Fed Proc 42:* 2566, 1983.
Hoogwerf B *et al*: *Clin Toxicol 14:* 47, 1979.
James SH, Schnoll SH: *Ibid 9:* 573, 1976.
Peterson RC, Stillman RC, eds: *NIDA Research Monograph 21:* August, 1978.

Treatment of Acute Drug Overdosage

Bledsoe BE *et al*, eds: *Prehospital Emergency Pharmacology*, Prentice-Hall, 1988.
Kline NS *et al*, eds: *Psychotropic Drugs: A Manual for Emergency Management of Overdose*, Medical Economics, Oradell, NJ, 1974.
Veltri JC, Temple AR: *Clin Toxicol 9:* 407, 1976.

CHAPTER 75

Preformulation

Louis J Ravin, PhD
Department of Pharmaceutics
Research & Development
Smith Kline & French Laboratories
King of Prussia, PA 19406

Galen W Radebaugh, PhD
Director of Pharmaceutics
Parke-Davis Pharmaceutical Research Division
Warner-Lambert Co
Morris Plains, NJ 07950

The attention presently being given to multisource pharmaceutical products regarding their equivalency places much emphasis on the formulation of these products. In some instances, the bioavailability of a drug formulation represents a quality parameter of enormous proportion. It is a matter of record that with certain drugs, depending on the formulation, the rate at which the drug substance becomes available can vary significantly from very high to none at all. As a result, the effectiveness of these formulations will range dramatically from that expected to no effect. Unfortunately, most examples are less dramatic and fall somewhere in between. The difference in the bioavailability of these drug products is less readily discernible, but nonetheless real. This has led to a great deal of confusion and information which, though understood by the scientist, is unclear and jumbled to the practitioner. That information which is available also has been interpreted differently by different individuals or groups, depending very often on the motivation, viewpoint and attitude of the interpreter.

Drug products indeed do vary in their bioavailability characteristics and this variation, in most instances, is related directly to formulation considerations. To optimize the performance of drug products, it is necessary to have a complete understanding of the physical-chemical properties of drug substances prior to formulating them into drug products. The development of an optimum formulation is not an easy task, and many factors readily influence formulation properties. Drug substances rarely are administered as chemical entities, but almost always are given in some kind of formulation. These may vary from a simple solution to a very complex drug delivery system. The complexity usually is not intentional, but rather is determined by the properties that are expected from or built into the dosage form and by the resulting composition that is required to achieve these qualities.

The high degree of uniformity, physiological availability and therapeutic quality expected of modern medicinal products usually are the results of considerable effort and expertise on the part of the formulating pharmacist. These qualities are attained by careful selection and control of the quality of the various ingredients employed, appropriate manufacturing according to well-defined processes and, most importantly, adequate consideration of the many variables that may influence the composition, stability and utility of the product. In dealing with the formulation of new products it has become necessary to apply the best research methods and tools in order to develop, produce and control the potent, stable and effective dosage forms which make up our modern medical armamentarium.

The pharmaceutical formulator has need for specialized areas of science in order to acquire scientific information about the drug substance which is necessary to develop an optimum dosage form. The pharmaceutical industry is in an era in which one can no longer rely on past experience to formulate. A thorough understanding of the physical and chemical properties as well as the pharmacokinetic and biopharmaceutical behavior of each drug substance being developed is necessary. In short, as much information as possible must be acquired about the drug substance very early in its development. This requires an interdisciplinary approach at the preformulation stage of development. Fig 75-1 schematically indicates that the development of any drug product requires a multidisciplinary approach, involving basic science, during the preformulation stage followed by applied science during the development stage.

This chapter will discuss the physical-chemical evaluation that takes place during the preformulation stage of development. In addition, consideration will be given to some specialized formulation ingredients that may require discretion in their selection.

Preformulation may be described as a stage of development during which the physical pharmacist characterizes the physical-chemical properties of the drug substance in question which are considered important in the formulation of a stable, effective and safe dosage form. Such parameters as crystal size and shape, pH-solubility profile, pH-stability profile, polymorphism, partitioning effect, drug permeability and dissolution behavior are evaluated. During this evaluation possible interactions with various inert ingredients intended for use in the final dosage form also are considered. The data obtained from this evaluation are integrated with data obtained from the preliminary pharmacologic and biochemical studies and provide the formulating pharmacist with information that permits selection of the optimum dosage form containing the most desirable inert ingredients for use in its development.

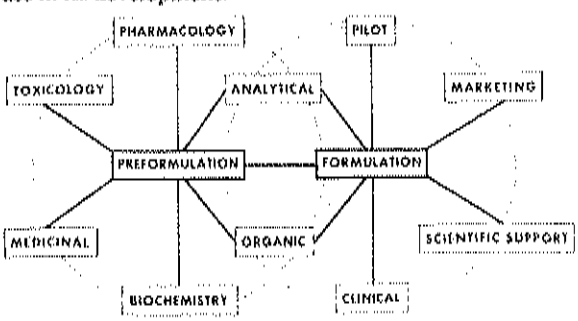


Fig 75-1. The wheels of product development.

Preformulation work usually is initiated after a compound has shown sufficient activity to merit further testing in humans. When this decision is made, the various disciplines begin to generate data essential for properly evaluating the performance of the drug substance. A stability-indicating analytical assay is very important. Since this often takes considerable time, it sometimes is necessary to rely on thin-layer chromatographic procedures to determine if a drug molecule is degrading. Accelerated testing procedures are used to promote breakdown of the compound being tested. Attempts are made to isolate and characterize the breakdown products in order to identify the mechanism of breakdown. This information provides a lead to the development pharmacist in his efforts to formulate the product.

During a preformulation study it is necessary to maintain some degree of flexibility. Problem areas must be identified early. For example, selection of a suitable salt form of the drug may be critical. Toxicity studies usually are scheduled early. Consequently, if the salt form under consideration has some deficiencies, they should be pointed out so that alternate salts may be prepared and evaluated prior to beginning toxicity studies.

When preformulation studies are initiated, the chemical usually is in short supply; 25 g of chemical is an ample supply, but many preliminary evaluations have been done with less. The initial supply usually originates as excess from batches prepared by the medicinal chemists. They usually have preliminary data such as melting point, solubility, spectral data and structure of the compound. The direction taken for the evaluation is determined by the structure and the intended dosage forms to be developed (eg, one would not waste time determining the stability of a solution of a compound if there was no interest in a liquid dosage form). Many areas must be evaluated critically for each compound, and it is essential that problem areas be identified early, otherwise delays could occur if a problem surfaced during the development phase for the compound. Some consequences of poor preformulation work are

Possible use of unsatisfactory salt form.
Poor stability of the active ingredient.
Testing compound of marginal activity.
Increased development costs.
Increased development time.

When preformulation studies are completed, the data are compiled and transferred to the development pharmacist, who, in turn, uses this information to plan his development work on the finished dosage forms.

Physical Properties

Description

Since the pure drug entity is in short supply at the outset of most preliminary evaluations, it is extremely important to note the general appearance, color and odor of the compound. These characteristics provide a basis for comparison with future lots. During the preparation of scale-up lots the chemist usually refines or alters the original chemical synthetic route. This sometimes results in a change in some of the physical properties. When this takes place, comparisons can be made with earlier lots and decisions made regarding solvents for recrystallization.

Taste usually warrants some consideration, especially if the drug is intended for oral use in pediatric dosage forms. In such cases consideration should be given to the preparation of alternate salt forms or possible evaluation of excipients that mask the undesirable taste.

Microscopic Examination

Each lot of drug substance, regardless of size, is examined microscopically and a photomicrograph taken. The microscopic examination gives a gross indication of particle size and characteristic crystal properties. These photomicrographs are useful in determining the consistency of particle size and crystal habit from batch to batch, especially during the early periods of chemical synthesis; if a synthetic step is changed, they also give an indication of any effect the change may have on crystal habit. One must keep in mind that the photomicrograph only gives a qualitative indication of particle size distribution; it always is necessary to do a particle-size analysis for a more accurate picture of the distribution of particles in any particular batch of drug substance.

Particle Size

The uses of pharmaceutical products in a finely divided form are diverse. From knowledge of their particle size, such drugs as griseofulvin, nitrofurantoin, spironolactone, procaine penicillin and phenobarbital have been formulated so as to optimize activity. Other drugs, formulated in suspension or emulsion systems, in inhalation aerosols or in oral dosage forms, may contain finely divided material as an essential component. One of the basic physical properties common to all these finely divided substances is the particle-size distribution, ie, the frequency of occurrence of particles of every size. What is of practical interest usually is not the characteristics of single particles but rather the mean characteristics of a large number of particles. It must be emphasized, however, that knowledge of size characteristics is of no value unless adequate correlation has been established with functional properties of specific interest in the drug formulation. Many investigations demonstrating the significance of particle size are reported in the literature. It has been shown that dissolution rate, absorption rate, content uniformity, color, taste, texture and stability depend to varying degrees on particle size and distribution. In preformulation work it is important that the significance of particle size in relation to formulation be established early. Preliminary physical observations sometimes can detect subtle differences in color. If this can be attributed to differences in particle-size distribution, it is important to define this distribution and recommend that more attention be given to particle size in preparing future batches of drug substance. This effect also is evident when preparing suspensions of poorly soluble materials. One may observe batch-to-batch differences in the color of a suspension which can be related to differences in particle size. Sometimes, when small particles tend to agglomerate, a subtle change in color or texture may be evident.

Sedimentation and flocculation rates in suspensions are in part governed by particle size. In concentrated deflocculated suspensions the larger particles exhibit hindered settling and the smaller particles settle more rapidly. In flocculated suspensions the particles are linked together into flocs which settle according to the size of the floc and porosity of the aggregated mass. Flocculated suspensions are preferred since they have less tendency to cake and are more rapidly dispersible. Thus, it is apparent that the ultimate height, H_u, of sediment as a suspension settles depends on particle size. The ratio H_u/H_o, or the degree of suspendibility as affected by particle size, is valuable information for the formulator in order to prepare a satisfactory dosage form.

The rate of dissolution of small particles usually is faster than that of larger ones because rate of dissolution depends on the specific surface area in contact with the liquid medium. This usually is described by the modified Noyes-Whitney equation for dissolution rate dA/dt

$$\frac{dA}{dt} = KS(C_s - C) \qquad (1)$$

where A is the amount of drug in solution, K is the intrinsic dissolution rate constant, S is the surface area, C_s is the

concentration of a saturated solution of the drug and C is the drug concentration at time t. The surface area of an object, regardless of shape, varies inversely with its diameter and confirms the above effect of particle size on dissolution rate. Solubility also has been observed to depend on particle size. Dittert, *et al*,[1] reported data for an experimental drug, 4-acetamidophenyl 2,2,2-trichloroethyl carbonate, which demonstrated that the dissolution rate and, in turn, bioavailability were affected by particle size. Although the ultimate amount of drug in solution may not be significant with respect to the dose administered, the formulator should be aware of this potential. With poorly soluble drugs it is extremely important to take these factors into account during the design of the dosage form.

Flow properties of drugs can be influenced by particle size, and particle size reduction to extremely small sizes (less than 10 μm) may be inadvisable for some drug substances. Entrapped air adsorbed on the surface of the particles and/or surface electrical charges sometimes impart undesirable properties to the drug. For example, adsorbed air at the drug-particle surface may prevent wetting of the drug by surrounding fluid, and electrically induced agglomeration of fine particles may decrease exposure of the drug surface to surrounding dissolution medium. Such effects act as dissolution rate-limiting steps since they minimize maximum drug surface-liquid contact.

Crystal growth is also a function of particle size. Finer particles tend to dissolve and subsequently recrystallize and adhere to larger particles. This phenomenon is referred to as *Ostwald ripening*. Protective colloid systems can be used to suppress this nucleation. Preformulators can generate information concerning the effectiveness of different colloids that is extremely important to the formulator when he is given the task of preparing a suspension dosage form.

Particle-size reduction may be deleterious for some drug substances. Increasing surface area by milling or other methods may lead to rapid degradation of a compound. Drug substances also may undergo polymorphic transformation during the milling process. The preformulator must always be cognizant of these potential problems, and whenever the decision is made to reduce particle size, the conditions must be controlled and the stability profile evaluated. If a problem does arise, it is the responsibility of the preformulator to note it and attempt to resolve it prior to turning the drug substance over to the formulating pharmacist.

Gastrointestinal absorption of a poorly soluble drug may be affected by the particle-size distribution. If the dissolution rate of the drug is less than the diffusion rate to the site of absorption and the absorption rate itself, then the particle size of the drug is of great importance. Smaller particles should increase dissolution rate and, thus, bring about more rapid gastrointestinal absorption. One of the first observations of this phenomenon was made with sulfadiazine. Blood-level determinations showed that the drug in suspension containing particles 1 to 3 μm in size was absorbed more rapidly and more efficiently than from a suspension containing particles 7 times larger. Maximum blood levels were about 40% higher and occurred 2 hours earlier. Increased bioavailability with particle-size reduction also has been observed with griseofulvin. The extent of absorption of an oral dose increased 2.5 times when the surface area was increased approximately sixfold. Micronized griseofulvin permits a 50% decrease in dosage to obtain a satisfactory clinical response.

On the other hand, it was found that with nitrofurantoin there was an optimal average particle size that minimized side effects without affecting therapeutic response. In fact, a commercial product containing large particles is available. For chloramphenicol, particle size has virtually no effect on total absorption but it significantly affects the rate of appearance of peak blood levels of the drug. After administration of 50-μm particles, as well as 200-μm particles, peak levels occurred in 1 hour; with 400-μm particles peak levels occurred in 2 hours; with 800-μm particles peak levels occurred in 3 hours. All four preparations had the same physiological availability, which implies that the absorption of chloramphenicol occurs uniformly over a major portion of the intestinal tract.

Reduction of particle size also may create adverse responses. For example, fine particles of the prodrug trichloroethyl carbonate were more toxic in mice than regular and coarse particles.[2] Increasing the surface area for water-soluble drugs, and possibly for weakly basic drugs, appears to be of little value. Absorption of weak bases usually is rate-limited by stomach emptying time rather than by dissolution. As previously mentioned, particle size is of importance only when the absorption process is rate-limited by the dissolution rate in gastrointestinal fluids.

The previous discussion considered the effect of particle size of the drug substance and its relationship to formulation. The particle size of the inert ingredients merits some attention. When one is concerned with particle size, all ingredients used in preparing the dosage form should be evaluated and some recommendation regarding their control should be made prior to full-scale development of a dosage form. It is recommended highly that particle size and its distribution be determined, optimized, monitored and controlled when applicable, particularly during early preformulation studies when the decision is made with regard to a suitable dosage form. The more common methods of determining particle size of powders used in the pharmaceutical industry include sieving, microscopy, sedimentation and stream scanning.

Sieving or Screening—Sieving or screening is probably one of the oldest methods of sizing particles and still is used commonly to determine the size distribution of powders in the size range of 325 mesh (44 μm) and greater. These data serve usually as a rough guideline in evaluating raw materials with regard to the need for milling. The basic disadvantages of screen analysis are the large sample size required and the tendency for blinding of the screens due to static charge or mechanical clogging. The advantages include simplicity, low cost and little skill requirement of the operator.

Microscopy—Microscopy is the most universally accepted and direct method of determining particle-size distribution of powders in the subsieve range, but this method is tedious and time-consuming. The preparation of the slide for counting particles is important because the sample must represent the particle-size distribution of the bulk sample. Extreme care must be taken in obtaining a truly representative sample from the bulk chemical. The cone and quartering technique usually gives a satisfactory sample. The sample should be properly suspended, dispersed and mixed thoroughly in a liquid which has a different refractive index from the particles being counted. A representative sample is mounted on a slide having a calibrated grid. For counting, random fields are selected on the slide and the particles are sized and counted. Between 500 and 1000 particles should be counted to make statistical treatment of the data meaningful.

Sedimentation—Sedimentation techniques utilize the dependence of velocity of fall of particles on their size. Application is made of the Stokes equation (see page 295) which describes a relationship between the rate at which a particle settles in a fluid medium to the size of that particle. Although the equation is based on spherical-shaped particles, it is used widely to determine the weight-size distribution of irregularly shaped particles. Data obtained by this procedure are usually reliable; however, the result may not agree with those obtained by other methods because of the limitations of the shape factor.

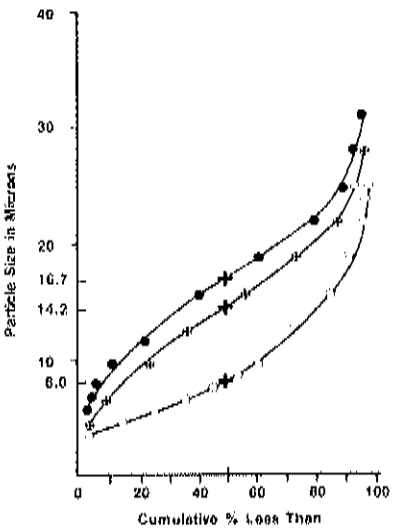


Fig 75-2. Particle size distribution of NBS glass beads (Standard Reference Material No 1003) expressed in terms of ○ = number of particles; ● = weight of particles; ⊗ = surface area of particles.

The *Andreasen Pipette Method* is used most commonly for sedimentation studies. Exact volumes are withdrawn at prescribed times and at a specified liquid depth. The liquid is evaporated and the residue of powder is weighed. The data are used in the Stokes equation and a weight-size distribution is calculated. Precautions must be observed with this method. Proper dispersion, consistent sampling, temperature control of the suspending medium and concentration should be achieved in order to obtain consistent results.

Stream Scanning—Stream scanning is a technique in which a fluid suspension passes through a sensing zone where the individual particles are electronically sized, counted and tabulated. The great advantage of this technique is that data can be generated in relatively short periods of time with reasonable accuracy. Literally thousands of particles can be counted in seconds and used in determining the size-distribution curve. The data are in a number of particles per class interval and can be expressed mathematically as the arithmetic mean diameter and graphed accordingly. Fig 75-2 illustrates a plot of typical data obtained for NBS Standard Reference Material No 1003.

The *Coulter Counter* and the *HIAC Counter* are used widely in the field of particle-size analysis in the pharmaceutical industry. They can be used to follow crystal growth in suspensions very effectively. Figure 75-3 shows the change in particle size with time for an aqueous suspension of Form I of an experimental drug. It appears that the growth of the particles decreases significantly after 6 hours. The photomicrograph shown in Fig 75-4 depicts the significant increase in particle size after 6 hours. Further treatment of the data as shown in Fig 75-5 enables one to establish rates of growth for suspended particles. Simply reading off the intercepts at the 1%, 2% or 3% oversize and plotting this increase in diameter with time enables one to calculate the rate of growth of particles in a suspension. This is shown in Fig 75-6.

Light Scattering—Light-scattering methods are generally fast, inexpensive and induce minimal artifacts. In general, such methods operate by measuring light diffraction from suspended particles without forming an image of the particles onto a detector. A typical unit is the laser diffraction particle sizer (*Malvern*). In it, a liquid dispersion of particles flows through a beam of laser light. Light scattered by the particles and the unscattered remainder are

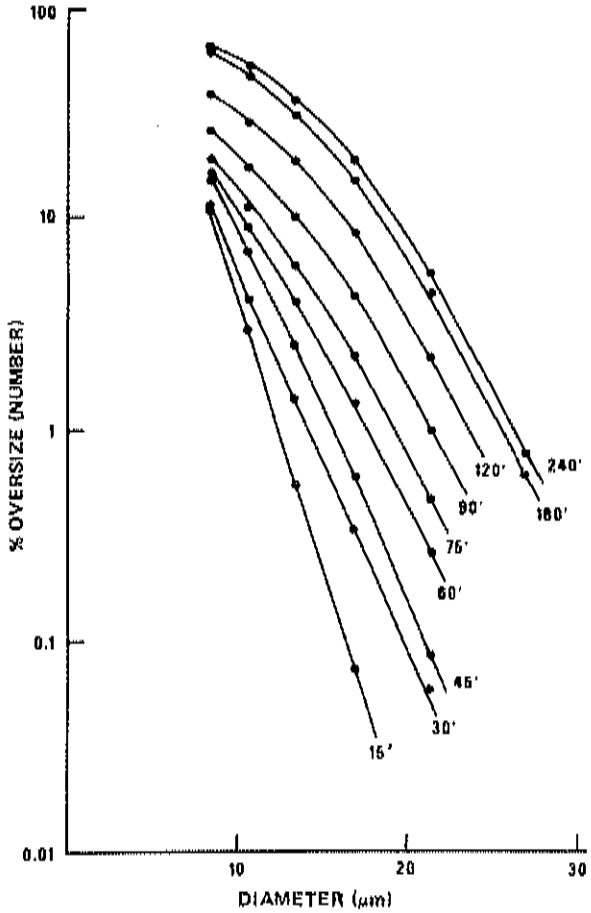


Fig 75-3. Change in particle size with time for an aqueous suspension of Form I of an experimental drug.

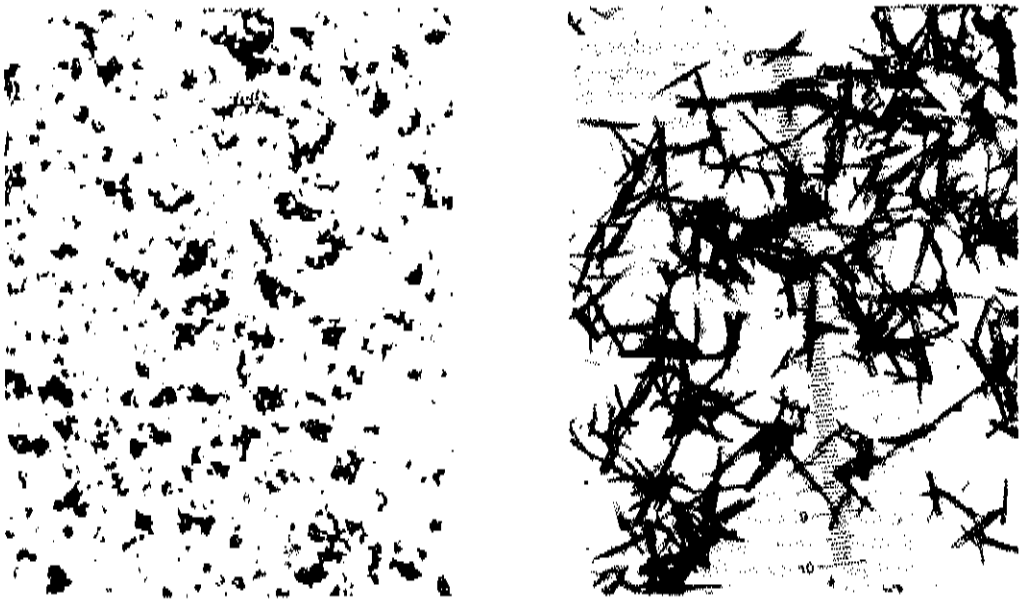
incident onto a receiver lens that forms a diffraction pattern of the scattered light. The scattered light and unscattered light then are gathered on detectors so the total light power is monitored as it allows the sample volume concentration to be determined. Each particle scatters light at a favored scattering angle that is related to its diameter. The detector provides an electronic output that makes it possible for a computer to deduce the volume-size distribution that gives rise to the observed scattering characteristics. Results may also be transformed to the equivalent surface or number distribution. Refer to Chapters 19 and 30.

Partitioning Effect

If an excess of liquid or solid is added to a mixture of two immiscible liquids, it will distribute itself between the two phases so that each becomes saturated. If the substance is added to the immiscible solvents in an amount insufficient to saturate the solutions, it still will distribute between the two layers in a definite concentration ratio. If C_1 and C_2 are the equilibrium concentrations of the substance in Solvent 1 and Solvent 2, the equilibrium expression becomes

$$\frac{C_1}{C_2} = k \qquad (2)$$

The equilibrium constant k is known as the distribution ratio or partition coefficient. Biologically, in order for a pharmacological response to occur, it is necessary that the drug molecule cross a biological membrane. The membrane, consisting of protein and lipid material, acts as a



FORM I

INITIAL SUSPENSION

FORM I

SUSPENSION AFTER 6 HOURS.

Fig 75-4. Photomicrographs showing change in crystal size for a suspension of Form I of an experimental drug.

lipophilic barrier to most drugs. The resistance of this barrier to drug transfer is related to the lipophilic nature of the molecule involved. (See Chapter 35.)

Understanding the partitioning effect and the dissociation constant enables one to estimate the site of absorption of a new chemical entity. If one assumes the stomach to have a pH range of 1.0 to 3.0 and the small intestines to have a pH range from 5 to 8, in most cases acidic drugs (pK_a 3) will be absorbed more rapidly in the stomach while more basic drugs (pK_a 8) will be absorbed more rapidly in the intestinal tract. There are exceptions, however. Some compounds have low partition coefficients and/or are ionized highly over the entire physiological pH range, but still show good bioavailability.

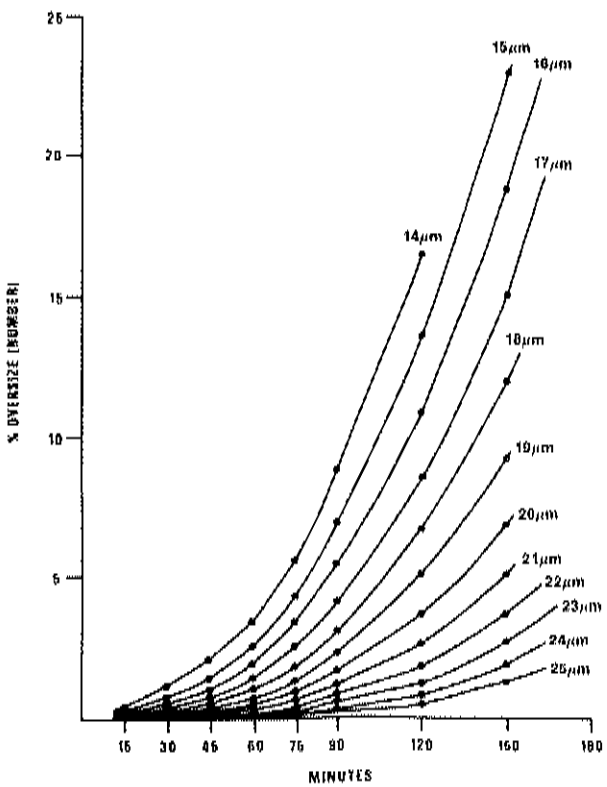


Fig 75-5. Change in cumulative count with time for an aqueous suspension of Form I of an experimental drug.

Polymorphism

A polymorph is a solid crystalline phase of a given compound resulting from the possibility of at least two different

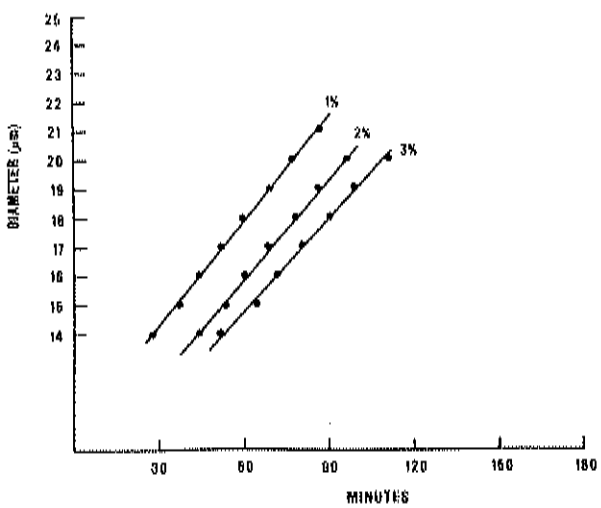


Fig 75-6. Rate of growth of Form I of experimental drug in aqueous suspension.

arrangements of the molecules of the compound in the solid state. The molecule itself may be of different shape in the two polymorphs, but that is not necessary and, indeed, certain changes in shape involve formation of different molecules and, hence, do not constitute polymorphism. Geometric isomers or tautomers, even though interconvertible and reversibly so, cannot be called polymorphs although they may behave in a confusingly similar manner.

A safe criterion for classification of a system as polymorphic is the following: two polymorphs will be different in crystal structure but identical in the liquid or vapor states. Dynamic isomers will melt at different temperatures, as do polymorphs, but will give melts of different composition. In time, each of these melts changes to an equilibrium mixture of the two isomers with temperature-dependent compositions. Some reported cases of polymorphism are undoubtedly dynamic isomerism, since the two behave quite similarly.

Polymorphism is the ability of any element or compound to crystallize as more than one distinct crystalline species, eg, carbon as a cubic diamond or hexagonal graphite. Different polymorphs of a given compound are, in general, as different in structure and properties as the crystals of two different compounds. Solubility, melting point, density, hardness, crystal shape, optical and electrical properties, vapor pressure, stability, etc all vary with the polymorphic form. In general, it should be possible to obtain different crystalline forms of a drug substance exhibiting polymorphism and, thus, modify the performance properties for that compound. To do so requires a knowledge of the behavior of polymorphs. There are numerous reviews on the subject of polymorphism. In addition, numerous indications of the importance of polymorphism in pharmaceuticals are reported in the literature. Extensive studies of polymorphism have been conducted on steroids, barbiturates, antihistamines and sulfonamides. Preformulation usually includes rigorous studies to determine the presence of polymorphs in new drug substances being prepared for preliminary investigation in test animals. Some of the parameters routinely investigated are the number of polymorphs that exist, relative degree of stability of the various polymorphs, presence of a glassy state, stabilization of metastable forms, temperature stability ranges for each polymorph, solubilities, method of preparation of each form, effect of micronization or tableting and interaction with formulation ingredients.

The initial task of the preformulator is to determine whether or not the drug substance being evaluated exists in more than one crystalline form. The following procedures are usually followed to cause crystallization of a metastable form:[3]

1. Melt completely a small amount of the compound on a slide and observe the solidification between crossed polars. If, after spontaneous freezing, a transformation occurs spontaneously or can be induced by seeding or scratching, the compound probably exists in at least two polymorphic forms. It is essential to prevent nucleation of the stable form by inducing supercooling. Supercooling can be induced by using a small sample size, holding the melt for approximately 30 sec about 10° above the melting point; carefully setting aside the compound without physical shock before observing it and rapid cooling of the compound.
2. Heat a sample of the compound on a hot stage and observe whether a solid-solid transformation occurs during heating.
3. Sublime a small amount of the compound and attempt to induce a transformation between the sublimate and the original sample by mixing the two in a drop of saturated solution of one of them. If the two are polymorphs, the more stable one will be more insoluble and will grow at the expense of the more soluble metastable form. This process will continue until the metastable form is transformed completely to the stable form. If the samples are not polymorphs, one may dissolve but the other will not grow. If the two are identical forms, nothing will occur.
4. Maintain an excess of the compound in a small amount of solvent held near the melting point of the compound. Isolate the suspended solid. Care should be taken to maintain the temperature during this step. Test the isolated material with an original sample using the procedure outlined in 3, above.
5. Recrystallize the compound from solution by shock-cooling, and observe a portion of the precipitated material suspended in a drop of the mother liquor. The drop then may be seeded with the original compound to check for solution-phase transformation. If the precipitate is a different polymorph, a solution-phase transformation should take place.

Once it has been established that polymorphism occurs, there are procedures which enable the preformulator to prepare the various forms in larger quantities for further evaluation and suitability for incorporation into dosage forms.

Once a compound has been shown to exist in more than one crystalline form, a number of techniques are available to identify the different polymorphic phases present. Each of these techniques could be successful in identifying the phase, but a combination of methods provides a means for isolation and identification of each crystalline modification. In order to confirm the presence of more than one crystalline form of a compound, it is advisable to identify the modifications present by more than one method. Using only one method for confirming the presence of polymorphs sometimes may be misleading.

Microscopy—Optical crystallography is used in the identification of polymorphs. Crystals exist in isotropic and anisotropic forms. When isotropic crystals are examined, the velocity of light is the same in all directions, while anisotropic crystals have two or three different light velocities or refractive indices. This method requires the services of a trained crystallographer.

Hot-Stage Methods—The polarizing microscope, fitted with a hot or cold stage, is very useful for investigating polymorphs. An experienced microscopist can tell quickly whether polymorphs exist; the degree of stability of the metastable forms; transition temperatures and melting points; rates of transition under various thermal and physical conditions and whether to pursue polymorphism as a route to an improved dosage form.

X-Ray Powder Diffraction—Crystalline materials in powder form give characteristic X-ray diffraction patterns made up of peaks in certain positions and varying intensities. Each powder pattern of the crystal lattice is characteristic for a given polymorph. This method has the advantage over other identification techniques in that the sample is examined as presented. Some care should be exercised in reducing and maintaining particle-size control. A very small sample size is needed and the method is nondestructive. This method has been used by several investigators in identifying polymorphs in pharmaceuticals.

Infrared Spectroscopy—This procedure is useful in identification of polymorphs. Solid samples must be used since polymorphs of a compound have identical spectra in solution. The technique can be used for both qualitative and quantitative identification.

Thermal Methods—Differential scanning calorimetry and differential thermal analysis have been used extensively to identify polymorphs. In both methods, the heat loss or gain resulting from physical or chemical transitions occurring in a sample is recorded as a function of temperature as the substance is heated at a uniform rate. Enthalpic changes, both endothermic and exothermic, are caused by phase transitions. For example, fusion, sublimation, solid-solid transition and water loss generally produce endothermic effects while crystallization produces exothermic effects. Thermal analysis enables one to calculate the thermodynamic parameters for the systems being evaluated. Heats of fusion can be obtained and the rate of conversion of polymorphs determined.

Dilatometry—Dilatometry measures the change in volume caused by thermal or chemical effects. Ravin and Higuchi[4] used dilatometry to follow the melting behavior of theobroma oil by measuring the specific volume of both rapidly and slowly cooled theobroma oil as a function of increasing temperature. The presence of the metastable form was shown by a contraction in the temperature range of 20° to 24°. This is illustrated in Fig 75-7. Dilatometry is extremely accurate; however, it is very tedious and time-consuming. It is not used widely.

Proton magnetic resonance, nuclear magnetic resonance and electron microscopy sometimes are used to study polymorphism.

Polymorphs can be classified into one of two types: (1) *enantiotropic*—one polymorphic form can be changed reversibly into another one by varying the temperature or pressure, eg, sulfur and (2) *monotropic*—one polymorphic form is unstable at all temperatures and pressures, eg, glyceryl stearates. At a specified temperature and pressure, only one polymorphic form will be thermodynamically stable. However, other metastable forms may exist under the

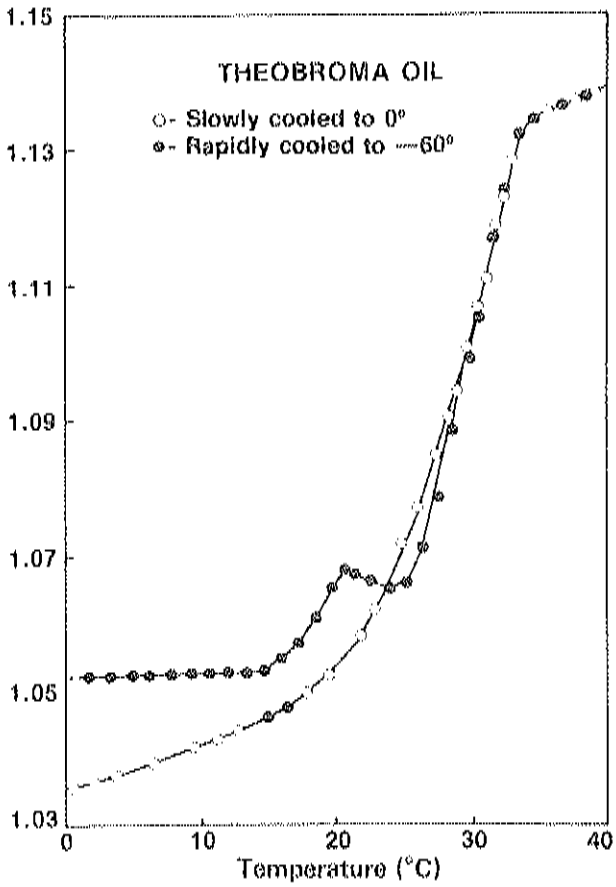


Fig 75-7. Dilatometric curves: theobroma oil, slowly and rapidly cooled.

same conditions. These metastable forms will convert to the stable lattice structures with time. The first indication of the significance of a polymorphic transformation in a pharmaceutical system was noted with novobiocin. The amorphous form of novobiocin was found to be well-absorbed; however, when formulated into a suspension, a reversion of the metastable form to the more stable crystalline form occurred resulting in poor absorption.

After it has been determined that a drug substance does exist in more than one crystalline form, the conditions under which each can be produced should be established. In this manner, proper crystallizing conditions can be maintained from batch to batch to ensure a uniform and acceptable raw material. Recrystallization solvent, rate of crystallization and other factors may cause one crystal form to dominate. During the preliminary investigation to establish these conditions, it is necessary to monitor the forms prepared. For example, during the preliminary work with an indole derivative, differential scanning calorimetry, X-ray analysis and infrared analysis were used to establish that polymorphs were present and that they could be prepared satisfactorily. Figs 75-8, 75-9, and 75-10 show the respective data for this conclusion. When polymorphs are shown to be present, experiments should be designed to determine whether or not the properties differ sufficiently to alter their pharmaceutic or biologic behavior.

Dissolution tests can be used initially to show differences in apparent equilibrium solubilities provided a discriminating solvent system is used. Fig 75-11 illustrates dissolution data for two polymorphs of an indole derivative which had similar dissolution in the medium used; however, when a more discriminating dissolution medium was used, it was

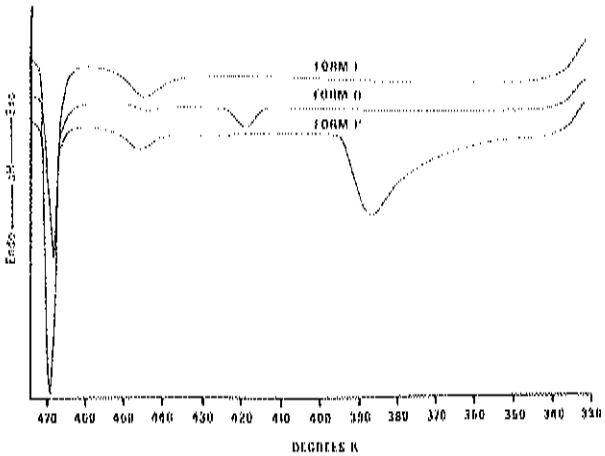


Fig 75-8. Thermograms for Forms I, I' and II of SK&F 30097.

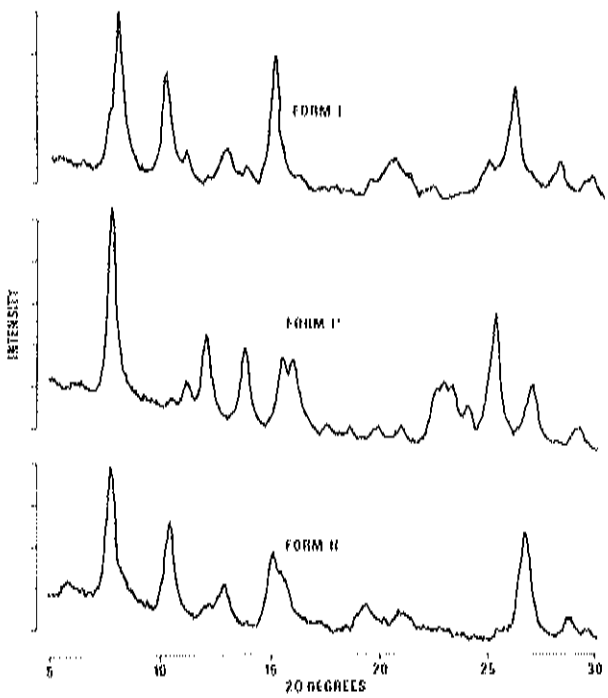


Fig 75-9. X-Ray diffractograms for Forms I, I' and II of SK&F 30097.

possible to show differences in their dissolution characteristics. This is illustrated in Fig 75-12. From the data presented for the indole derivative, it was concluded that there would be no appreciable difference in the availability of the two forms if they were to be administered orally in a solid dosage form. Subsequent testing in animals confirmed this. The Nernst equation relates the rate of concentration increase to the solubility of a dissolving solid and is commonly written as

$$\frac{dc}{dt} = \frac{AD}{Vh}(C_s - C_t) \tag{3}$$

where A is the area of the dissolving interface of the solid, D is the diffusion coefficient of the solute in the solvent, V is the volume of the solvent, h is the thickness of the diffusion layer and C_s and C_t are concentration of the solute at saturation and at time t, respectively. The equation reduces to

$$\frac{dc}{dt} = \frac{AD}{Vh} C_s \tag{4}$$

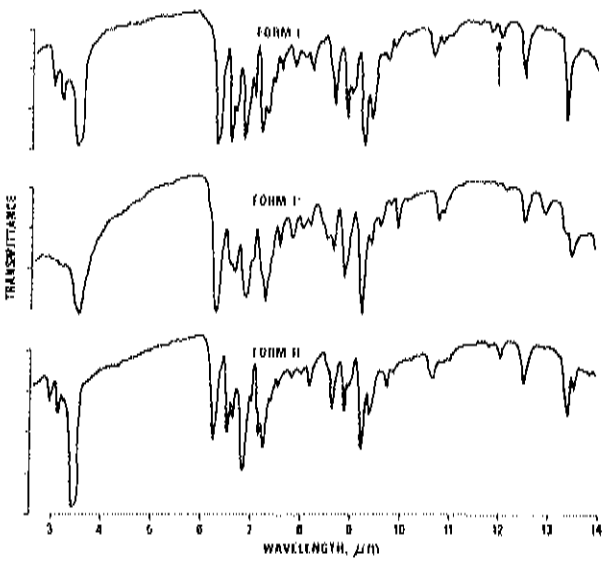


Fig 75-10. Infrared spectra of Forms I, I' and II of SK&F 30097.

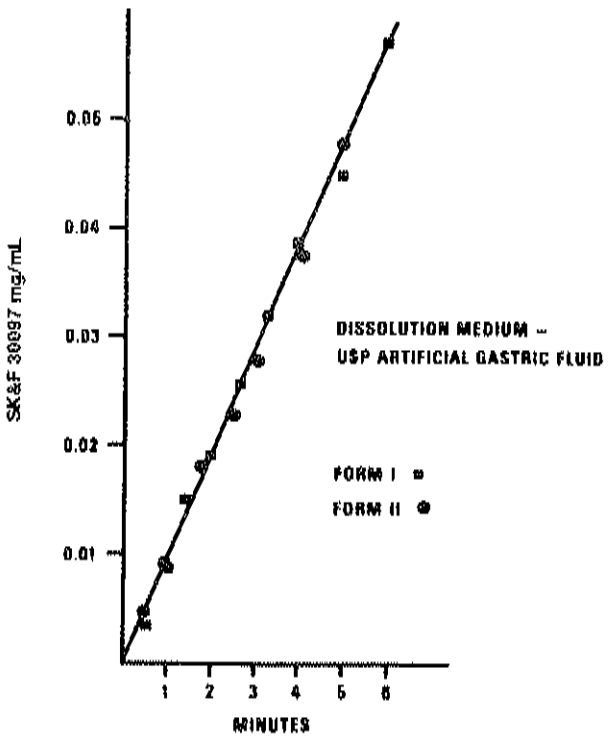


Fig 75-11. Dissolution behavior of Forms I and II of SK&F 30097 in artificial gastric fluid.

for the experimental conditions where $C_s > C_t$. Since D is a property of the solute molecule and the solvent, it is independent of the solid-state form. The experimental conditions can be selected such that A, V and h can be maintained constant in measuring the dissolution rates of different polymorphic forms. The dissolution rate then is directly proportional to C_s, the saturation solubility, and the differences in the solubilities, can be related to their free energies.

The solubility and dissolution behavior of several polymorphs of chloramphenicol palmitate have been determined. Figs 75-13 and 75-14 illustrate the data obtained at several temperatures. It is apparent from the dissolution behavior that the maximum values obtained were good ap-

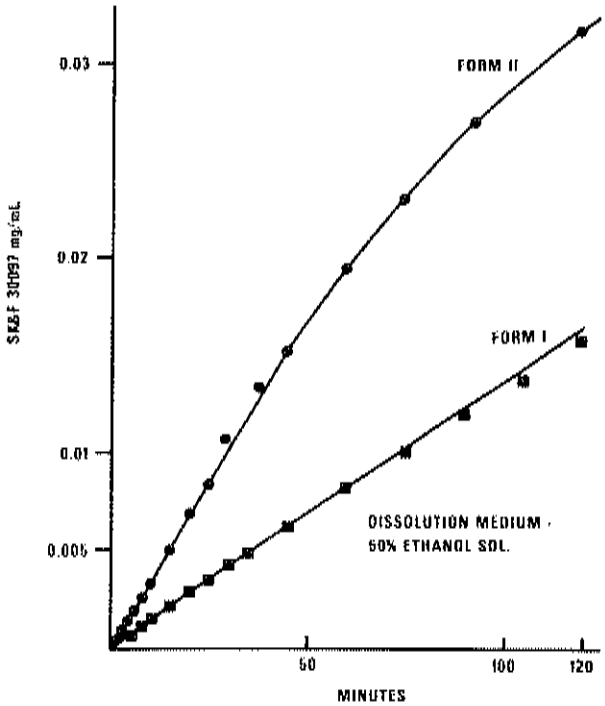


Fig 75-12. Dissolution behavior of Forms I and II of SK&F 30097 in 50% ethanol solution.

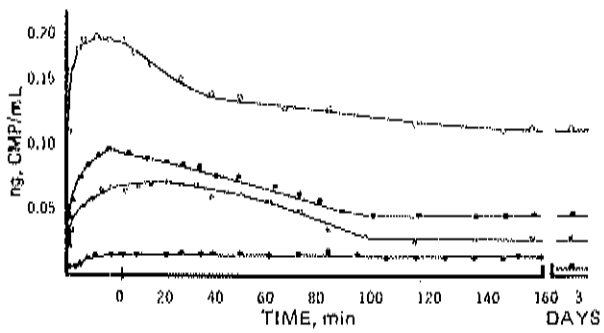


Fig 75-13. Dissolution curves for Polymorph C of chloramphenicol palmitate in 35% *t*-butyl alcohol and water at 30, 20, 15 and 6°. Key: 30°, O—O; 20°, ■---■; 15°, △—△; 6°, ●---●.

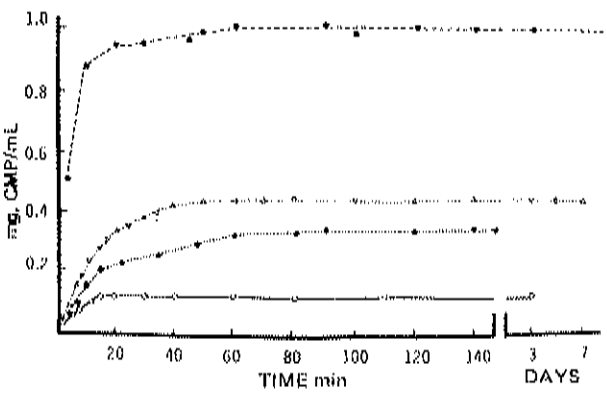


Fig 75-14. Dissolution curves for Polymorphs A and B of chloramphenicol palmitate in 35% *t*-butyl alcohol and water at 30 and 38°. Key: Polymorph A, 30°, O—O; Polymorph B, 30°, △. . .△; Polymorph A, 38°, ♦- - -♦; Polymorph B, 38° ●- - -●.

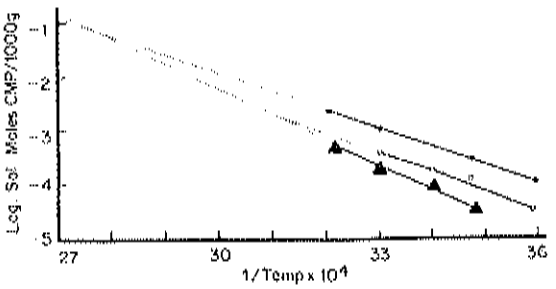


Fig 75-15. The van't Hoff type plot for Polymorphs A, B, and C of chloramphenicol palmitate. Key: Polymorphs A ▲; B ●—●; and C ○—○.

Table I—Thermodynamic Values Calculated for Polymorphs *A*, *B* and *C* of Chloramphenicol Palmitate[5]

Polymorph	Transition Temp. (°C) to Form *A*	Heat of Solution, kcal/mole	ΔG_t, cal/mole[a]	ΔS_{303} esu	ΔS_{trans} esu[a]
A	----	21.8	----	----	----
B	88	16.4	−774	−18	−17
C	50	17.2	−465	−13	−14

[a] Calculated for the conversion to Polymorph *A*.

proximations of the solubility of the various forms. Therefore, obtaining data at several temperatures would enable one to calculate the thermodynamic quantities involved in the transition from the metastable to the stable form. A plot of the solubility data as a function of temperature in a typical van't Hoff fashion is shown in Fig 75-15. The straight-line relationship enables one to calculate the heats of solution for the various forms and also, by extrapolation, to approximate the transition temperatures for the various forms. These values are shown in Table I.[5]

At constant temperature and pressure, the free-energy differences between the polymorphs can be calculated by

$$\Delta G_t = RT \ln \frac{C_s \text{ Polymorph A}}{C_s \text{ Polymorph B}} \qquad (5)$$

This equation relates the solubility, C_s, of the polymorphic forms at a particular temperature, T, to the free energy differences, ΔG_t. Table I also contains the free-energy differences calculated for the polymorphs. The enthalpy changes also can be determined for the various transitions by subtracting the heat of solution derived for the stable form from that of the metastable form. Also, at any particular temperature, T, the entropy for the transition of polymorphs can be evaluated by the following relationship

$$\Delta S_t = \frac{\Delta H_{B\to A} - \Delta G_t}{T} \qquad (6)$$

The values computed for the transitions also are included in Table I. At the transition temperature, ΔG_t is equal to zero and the entropy can be calculated, neglecting the free-energy term in Eq 6.

The thermodynamic relationships discussed are based on the assumption that Henry's Law is obeyed. Knowledge of these thermodynamic relationships enables the preformulator to select more rationally the more energetic polymorphic form of the drug being investigated for further pharmacological studies and also to have a preliminary assessment of its probable stability.

When a preformulation group inadequately investigates polymorphic drug forms, problems may develop during the development stage. Crystal growth in suspensions resulting in poor uniformity, poor appearance, poor bioavailability, transformation occurring during milling or granulation resulting in changes in the physical and biological characteristics, inadequate pharmacological response and poor chemical stability are typical problems that may become evident.

Solubility

In dealing with new drug substances, it is extremely important to know something about their solubility characteristics, especially in aqueous systems since they must possess some limited aqueous solubility to elicit a therapeutic response. When a drug substance has an aqueous solubility less than 1 mg/mL in the physiologic pH range (1–7), a potential bioavailability problem may exist and preformulation studies should be initiated to alleviate the problem. Equilibrium solubility of the drug substance should be determined in a solvent or solvent system which does not have any toxic effects on the test animal. This is done by placing an excess of drug in a vial with the solvent. The vial is agitated at constant temperature and the amount of drug determined periodically by analysis of the supernatant fluid. Equilibrium is not achieved until at least two successive samples have the same result. Experience with solubility determinations would indicate that equilibrium is usually attained by agitating overnight (approximately 24 hours). Solubility determinations can be conducted at several temperatures since the resultant drug products ultimately will be subjected to a wide variation in temperature.

If the solubility of the drug substance is less than the required concentration necessary for the recommended dose, steps must be taken to improve its solubility. The approach taken usually will depend on the chemical nature of the drug substance and the type of drug product desired. If the drug substance is acidic or basic, its solubility can be influenced by pH. Through the application of the Law of Mass Action, the solubility of weakly acidic and basic drug substances can be predicted as a function of pH with a considerable degree of accuracy, using the following equations for the weakly acidic and basic drugs.

$$\textit{Weak Acid} \qquad S_t = K_s\left(1 + \frac{K_a}{[H^+]}\right) \qquad \textit{Weak Base} \qquad S_t = K_s\left(1 + \frac{[H^+]}{K_a}\right) \qquad (7)$$

There are many drug substances for which pH adjustment does not provide an appropriate means for effecting solution. Very weakly acidic or basic drugs may require a pH that could fall outside the accepted tolerable physiological range or may cause stability problems with formulation ingredients. For example, an experimental indole had an equilibrium solubility at pH 1.2 of approximately 50 mg/mL. However, when the pH of this system was increased to approximately 2.0, the solubility decreased to less than 0.1 mg/mL. In cases like this one, or with nonelectrolytes, it is necessary to use some other means of achieving better solubility.

Cosolvent systems have been used quite effectively to achieve solubility for poorly soluble drug substances under investigation. Propylene glycol, glycerin, sorbitol and polyethylene glycols have enjoyed a wide range of success in this area. They have been very useful and generally acceptable for improving solubility. Additional solvents such as glyceryl formal, glycofurol, ethyl carbonate, ethyl lactate and dimethylacetamide have been cited in a review article by Spiegel and Noseworthy;[6] however, it must be emphasized that with the possible exception of dimethylacetamide all of these solvents have not been used in oral products and their

acceptability may be doubtful. The number of vehicles readily available to improve solubility is rather limited, yet the frequency of their use is rather high. Solubilizing a new drug substance can improve its availability. For example, when a triazinoindole was administered in a 0.02% solution it showed an equivalent response in antiviral activity to a 2.5% suspension. Information generated early in the preformulation stage can result in a refinement of the dosage regimen and allow for a more accurate estimation of the effective dose.

Cosolvents usually serve a twofold purpose in many pharmaceutical liquid products. They not only effect solution of the drug substance but also improve the solubility of flavoring constituents added to the product. Ideally, in determining the appropriate ratio of cosolvents to achieve the concentration one must achieve, it is recommended to effect solution at the concentration desired and then place the solution at 5° and allow it to equilibrate. If precipitation occurs under these conditions, it may be necessary to alter the cosolvent ratio.

The use of surfactants of various types—nonionic, cationic or anionic—as solubilizing agents for medicinal substances is widespread (see Chapter 19 for illustrations of specific uses). The effect of Triton WR-1339 in solubilizing several steroids is shown in Fig 75-16.[7] The effect of an anionic, a cationic and a nonionic surfactant on the solubility of an antianginal compound being considered for clinical trials is shown in Fig 75-17. From such data investigators may be guided in the selection of solubilizing agents for use in preparations to be studied in humans, but it must be emphasized that the acceptability of a particular solubilizing agent depends also on other factors that determine its suitability for the intended use. For example, surfactants are known to interact with some preservatives and thereby decrease preservative action, for which reason the preformulator should always recommend some type of biological test to demonstrate that the activity of the drug substance being studied is not reduced when it is solubilized by a surfactant.

Complexation phenomena sometimes can be used to impart better solubility characteristics. However, the degree of association and the extent to which solubility can be increased generally is not adequate for use in pharmaceutical products. In addition, many complexing agents have physiological activity. The most noteworthy example of the

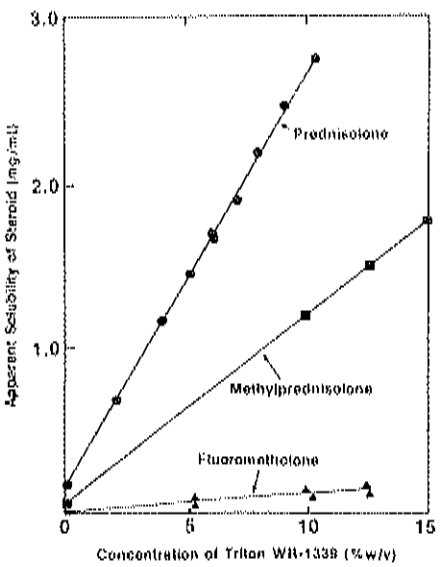


Fig 75-16. The effect of varying concentrations of Triton WR-1339 in water on the solubility of some anti-inflammatory steroids.[7]

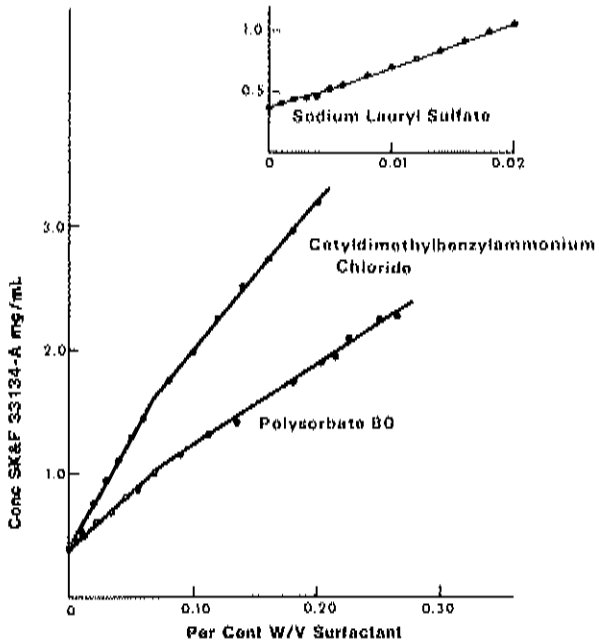


Fig 75-17. Effect of surfactant concentration on the solubility of SK&F 33134-A.

utility of complexation to enhance solubility is the PVP-iodine complex. Hydrotropy sometimes can be used to enhance solubility. High concentrations of urea, salicylates and xanthines have been used successfully on several occasions. Again, the concept is available but the increase in solubility normally observed is not adequate for use in pharmaceutical products.

Salt Formation

Salt-forming agents often are chosen empirically by the pharmaceutical chemist primarily on the basis of the cost of raw materials, ease of recrystallization and percentage yield. Unfortunately, there is no reliable way of predicting the influence of a particular salt species on the behavior of the parent compound in dosage forms. Furthermore, even when many salts of the basic compound have been prepared, there are no effective screening techniques which make the selection process of the salt an easier task for the pharmacist. The fundamental considerations which may have some influence on salt selection are physical and chemical stability, hygroscopicity, flowability and solubility.

The number of salt forms available to the chemist is large. Table II lists the cations and anions present in FDA-approved commercially marketed salts of pharmaceutical agents.[8] The monoprotic hydrochlorides have been the most frequent choice of the available anionic salt-forming radicals, while sodium has been the most predominant cation. During preformulation evaluation it is extremely important to establish that the particular salt form in question will have properties that will result in a minimum of problems during the development of the dosage forms. Since toxicity studies usually are initiated soon after a compound has been designated for further studies in man, it is important that the salt form selected has been given a critical evaluation to determine whether or not its properties are suitable.

Since physical and chemical stability are vital to any pharmaceutical product, it is imperative that the preformulator evaluate both parameters. A systematic determination of the thermal stability, solution stability (at several pH's) and

Table II—FDA-Approved Commercially Marketed Salts

Anion	Percent[a]	Anion	Percent[a]
Acetate	1.26	Iodide	2.02
Benzenesulfonate	0.25	Isethionate[f]	0.88
Benzoate	0.51	Lactate	0.76
Bicarbonate	0.13	Lactobionate	0.13
Bitartrate	0.63	Malate	0.13
Bromide	4.68	Maleate	3.03
Calcium edetate	0.25	Mandelate	0.38
Camsylate[b]	0.25	Mesylate	2.02
Carbonate	0.38	Methylbromide	0.76
Chloride	4.17	Methylnitrate	0.38
Citrate	3.03	Methylsulfate	0.88
Dihydrochloride	0.51	Mucate	0.13
Edetate	0.25	Napsylate	0.25
Edisylate[c]	0.38	Nitrate	0.64
Estolate[d]	0.13	Pamoate (Embonate)	1.01
Esylate[e]	0.13	Pantothenate	0.25
Fumarate	0.25	Phosphate/diphosphate	3.16
Gluceptate[g]	0.18	Polygalacturonate	0.13
Gluconate	0.51	Salicylate	0.88
Glutamate	0.25	Stearate	0.25
Glycollylarsanilate[h]	0.13	Subacetate	0.38
Hexylresorcinate	0.13	Succinate	0.38
Hydrabamine[i]	0.25	Sulfate	7.46
Hydrobromide	1.90	Tannate	0.88
Hydrochloride	42.98	Tartrate	3.54
Hydroxynaphthoate	0.25	Teoclate[j]	0.13
		Triethiodide	0.13

Cation	Percent[a]	Cation	Percent[a]
Organic:		Metallic:	
Benzathine[k]	0.66	Aluminum	0.66
Chloroprocaine	0.33	Calcium	10.49
Choline	0.33	Lithium	1.64
Diethanolamine	0.98	Magnesium	1.31
Ethylenediamine	0.66	Potassium	10.82
Meglumine[l]	2.29	Sodium	61.97
Procaine	0.66	Zinc	2.95

[a] Percent is based on total number of anionic or cationic salts in use through 1974. [b] Camphorsulfonate. [c] 1,2-Ethanedisulfonate. [d] Lauryl sulfate. [e] Ethanesulfonate. [f] Glucoheptonate. [g] *p*-Glycollamidophenylarsonate. [h] *N,N'*-Ditdehydroabietylethylenediamine. [i] 2-Hydroxyethanesulfonate. [j] 8-Chlorotheophyllinate. [k] *N,N'*-Dibenzylethylenediamine. [l] *N*-Methylglucamine.

light-sensitivity of the drug substance provides essential input toward the selection of the most suitable derivative. Studies usually are initiated early to identify problems. Samples of the salts in question usually are placed under exaggerated conditions of heat and light in the presence and absence of moisture and subsequently analyzed to determine the amount of breakdown. In many instances stability-indicating analytical methods may not be available. In these cases it is necessary to resort to thin-layer chromatography to establish a qualitative assessment of stability. At the same time, samples are placed under high-humidity conditions and weighed periodically to determine the degree of hygroscopicity of the compounds. Compounds that have a tendency to adsorb or absorb moisture may present flowability problems during encapsulation.

Solubility characteristics also are evaluated. When a particular salt form has very good solubility (greater than 10%) it sometimes is difficult to prepare a suitable granulation using an aqueous granulating fluid, especially for high doses. Granulations prepared by these methods will not dry satisfactorily or the granulation will not flow uniformly from the hopper, resulting in a large weight variation during the compression stage. A critical evaluation of this type with different salt forms has been proven quite effective in enabling the preformulator to make the selection of the salt form of choice for further development.

Compressibility and Compactibility

Tablets remain a preferred dosage form, and information obtained during preformulation studies on the ability of powdered drugs to be compressed and compacted can be a valuable aid to formulators. Compressibility and compactibility relate directly to tableting performance. Compressibility can be defined as the ability of a powder to decrease in volume under pressure, while compactibility can be defined as the ability of a powder to be compressed into a tablet of a certain strength or hardness. Even though powdered drugs usually are formulated with excipients to modify compression and compaction properties, the properties of the powdered drug alone may be the primary determinant of its ability to be manufactured into a tablet. Significant differences in compression and compaction behavior often can be observed in different lots of the same drug. For example, changes in crystallization or milling procedures may produce differences in behavior.

Compression and compaction most often are evaluated by measuring the tensile strength and hardness of compacts. Tensile strength commonly is measured by diametral compression of round tablets, where the analysis of strength accounts for the dimensions of the tablet.[9] Transverse compression of square compacts between platens narrower than the compact is reported to provide more reproducible results on a wider variety of powders.

Hardness can be defined as the resistance of a solid to local permanent deformation. Deformation hardness tests usually are measured by static impression or dynamic methods. The static method involves the formation of a permanent indentation on a solid surface by a gradual and regularly increasing stress load. Hardness is determined by the load and size of the indentation and is expressed as force per unit area. In dynamic tests, the solid surface is exposed to an abrupt impact such as a swinging pendulum or an indenter allowed to fall under gravity onto the surface. Hardness then is determined from the rebound height of the pendulum or the volume of the resulting indentation.

Hiestand has used adaptations of a compression test and a hardness test to obtain measurements that are used to formulate three dimensionless parameters or indices.[10] The indices are used to characterize the relative tableting performance of individual components or mixtures. The *Strain Index* is the ratio of dynamic indentation hardness to reduced Young's modulus. The *Bonding Index* is the ratio of tensile strength to indentation hardness. The *Brittle Fracture Index* is obtained by comparing the tensile strengths of square compacts with and without a hole at their center. The indices themselves do not measure intrinsic properties of a chemical compound, but rather the traits that influence the tableting performance of a specific lot of chemical. It is necessary to know the magnitude of all three indices to predict the variety of tableting properties that may be incurred. Such information can act as a guide in selecting excipients to overcome problem properties of a drug ingredient.

Chemical Properties

The evaluation of the physical and chemical stability of a new drug substance is an important function of the preformulation group. The initial work should be designed to identify those factors that may result in an alteration of the drug substance under study. The physical pharmacist initially can anticipate the possible type of breakdown that a compound will be subjected to by examination of the chemical structure of the compound. For example, esters and amides are sensitive to hydrolytic degradation while acri-

danes and catecholamines are sensitive to oxidative degradation. With this preliminary knowledge one may more effectively design studies to identify the problems early. At this point the primary concern is not the pathway or mechanism of degradation. A stability-indicating method of analysis usually is not available early in the preformulation phase. Techniques such as thin-layer chromatography, diffuse reflectance and thermal analysis can be used to provide data to assess preliminary stability. Sometimes, the preliminary evaluation is complicated by the presence of impurities. It is essential that the drug under study be pure before any stability tests are undertaken. The presence of impurities can lead to erroneous conclusions in the preformulation evaluation.

Drug Substance Stability—It is extremely important to determine the stability of the bulk chemical as early as possible. One hardly would expect to prepare stable dosage forms with a chemical substance that was not stable in the pure state. Samples of the chemical are subjected usually to various conditions of light, heat and moisture in the presence and absence of oxygen. The chemical is placed in sealed vials with and without moisture and stored at various elevated temperatures which may vary to some degree from laboratory to laboratory. Light-sensitivity is measured by exposing the surface of the compound to light. Sunlamps are sometimes used to exaggerate light conditions. Hygroscopicity is evaluated by placing the chemical in open petri dishes at relative humidities from 30 to 100%. The samples are monitored regularly for physical changes, moisture pick-up and chemical degradation.

Most drug substances are either stable at all conditions, stable under special conditions of handling, unstable with special handling or completely unstable. When drug substances are found to have some stability problems, it may be important to define the pathway of degradation and initiate studies to stabilize the compound with appropriate additives.

At this point, it may be advisable to consider some of the more prominent reactions accounting for instability of new drug substances. Obviously, some compounds will not undergo any appreciable decomposition if kept dry and away from air in a sealed container. It must always be assumed that the new drug substance is in some kind of formulation environment that may lead to instability problems.

Hydrolytic Degradation—Hydrolysis is probably the degradative process encountered most frequently in the formulation of new drugs. It is safe to assume that most new drugs will be exposed to water at some stage during processing or during storage; hence, hydrolysis may occur unless the conditions are optimum. Hydrolysis occurs with esters, amides, salts of weak acids and strong bases and thioesters, among others. A few drug compounds that undergo hydrolytic degradation are procaine, penicillin, aspirin and chlorothiazide.

From a kinetic standpoint, hydrolysis reactions are second-order reactions because the rate is proportional to the concentration of two reactants. However, in aqueous solutions, since water is usually present in excess and at relatively constant concentration, the reactions are treated experimentally as monomolecular or first-order reactions. This simplification permits calculations of the extent of decomposition under precise experimental conditions by less-complicated means. Extrapolation of the exaggerated rates to room temperature makes it possible to establish more expeditiously shelf-life stability of potential new drug products.

The rate of hydrolysis can be affected by temperature and by hydrogen or hydroxyl ion concentration when the hydrolytic process is dependent on pH. Fig 75-18 shows the pseudo-first-order behavior as a function of pH for carbuterol in aqueous solution at constant ionic strength at 85°. The

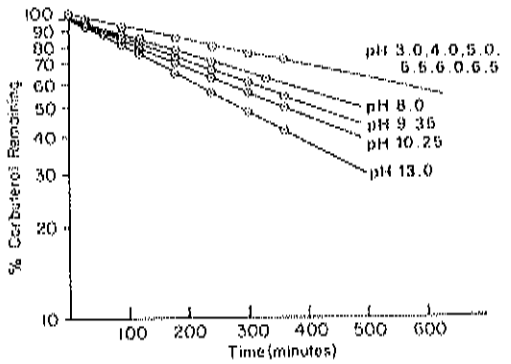


Fig 75-18. Effect of pH on carbuterol degradation at 85° ($\mu = 0.5$).

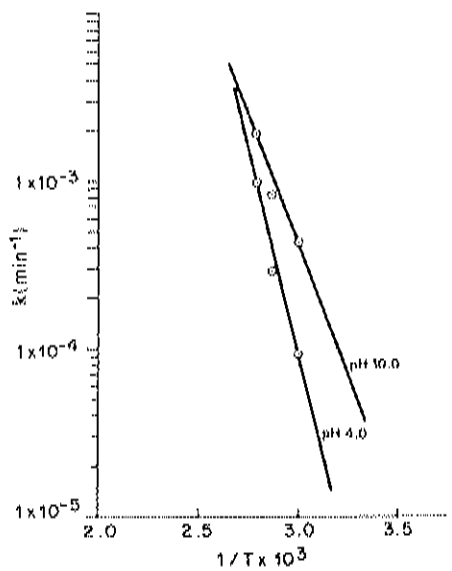


Fig 75-19. Typical Arrhenius-type plot depicting the temperature dependency of carbuterol hydrolysis at pH 4.0 and 10.0.

effect of temperature is illustrated in Fig 75-19 for carbuterol at pH 4.0 and 10.0 respectively.[11] For solids, the amount of moisture present is minimal. When considering a drug substance that undergoes hydrolytic degradation, studies are designed to establish the conditions of pH and buffer concentration where minimum decomposition occurs. There sometimes is a wide range of pH adjustment that a drug substance can tolerate. For example, idoxuridine was shown to have maximum stability over a pH range from 2.0 to 6.0. Fig 75-20 shows the pH-stability profile.[12] Another drug substance, carbuterol, hydrolyzed by an intramolecular process showed maximum stability over a wide pH range. Even though these compounds exhibited a wide range of pH for optimum stability in aqueous solution, they could not be formulated and provide products with satisfactory shelf lives without special cosolvent systems and/or special storage conditions. Cefazolin was shown to have a narrow pH range for maximum stability as indicated in Fig 75-21.[13] Buffering aqueous solutions to provide a pH for optimum stability can lead to stability problems. Stability sometimes is affected by buffer concentration; for example, carbuterol stability was shown to be affected by phosphate buffer concentration.

Another manner in which the physical pharmacist can overcome an instability due to hydrolysis is to recommend the preparation of an insoluble salt form or to prepare a solid

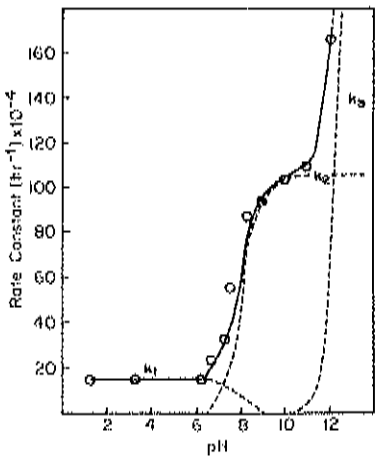


Fig 75-20. Plot showing pH-rate profile for hydrolysis of idoxuridine at 60°. Circles represent experimental results. Solid line corresponds to theoretical pH-rate profile. Broken line designates contribution of k_1, k_2 and k_3 at any pH value.

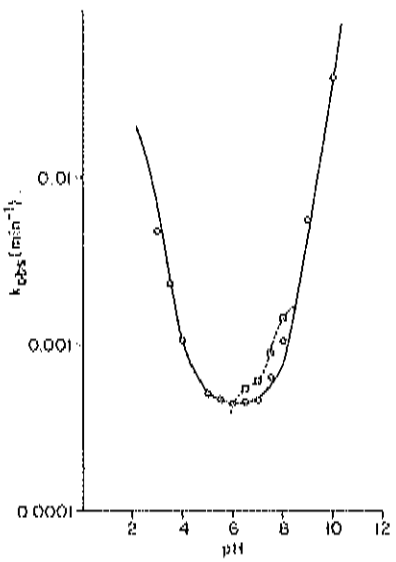


Fig 75-21. pH-Rate profile of cefazolin degradation in aqueous solution at 60° (μ = 0.5). Solid line: theoretical profile; circles: experimental profile; squares: rates uncorrected for buffer effect.

dosage form. Insoluble chlorothiazide is stable in neutral aqueous suspensions, but solutions of the sodium salt at relatively high pH decompose rapidly. Frequently, the replacement of water by some other solvent, such as alcohol or the polyhydroxy solvents, reduces the hydrolytic rate of degradation for some systems. Acetylsalicylic acid suspensions containing high concentrations of sorbitol improved stability. Ampicillin also was shown to be more stable when the concentration of alcohol was increased. The formation of molecular complexes with aromatic esters greatly reduces the hydrolytic rate of degradation.

It also has been shown that stability of some compounds may vary depending on whether or not they exist in the micellar or nonmicellar state. For example, a difference in the chemical stability of penicillin exists in the micellar state from that in the monomeric state.

Oxidation—Oxidative degradation is as important as hydrolysis in the preliminary stability evaluation of new-drug substances. Studies should be initiated to establish the oxidative route, then steps should be taken to determine what additives can minimize the degradation. Oxidative degradation is common with many drug compounds. Ascorbic acid, epinephrine, vitamin A, chlorpromazine, isoproterenol, morphine, resorcinol and unsaturated fats and oils are subject to oxidative degradation. The oxidation reaction depends on several factors, including temperature, oxygen concentration in the liquid, impurities present and the concentration of the oxidizable component. The temperature effect in solutions is usually minimal; however, in the dry state it is more pronounced since other factors such as moisture dictate its stability behavior.

Initially, it is important to establish that oxidation is taking place. Solutions of the drug substance in question are exposed to various exaggerated conditions of light and oxygen tension in amber and flint-glass containers. Samples are analyzed for degradation. When it has been established that the oxidative route is the principal pathway for degradation, appropriate additives are used to determine what effect they might have on the stability. Sometimes pH is critical, since a great number of oxidation-reduction processes depend on the concentration of hydrogen or hydroxyl ions. Light usually accelerates degradation, thus the storage of products in dark containers does much to preserve stability. Photochemical changes many times involve the formation of other reactive compounds or free radicals which function to propagate the decomposition, once started. Auto-oxidation may occur in the absence of light when susceptible materials, such as fats and oils, are stored in the presence of air. The auto-oxidation of phenolic compounds is of special significance since compounds such as epinephrine and isoproterenol degrade in this manner. Heavy metal ions, eg, cupric and ferric, accelerate the oxidation of ascorbic acid and the phenothiazines. Frequently, only trace quantities of these ions, occurring as impurities, may be sufficient to cause an increased rate of decomposition. This can be a consistent problem since many of the so-called inert ingredients may have heavy metal contaminants.

The oxygen concentration in solution is a factor in many cases and often depends upon the temperature of storage or the solvent employed. Oxygen is more soluble in water at lower temperatures so that oxygen-dependent reactions can sometimes proceed more rapidly at the lower temperatures. Ascorbic acid is more stable in 90% propylene glycol or in Syrup USP than in water, presumably because of the lower oxygen concentration in these vehicles. Oxidative degradation is an extremely complex process since the overall rate is dependent upon several factors. Preparations sensitive to oxidation are sometimes stabilized by effectively removing the oxygen and by the addition of suitable additives. Nitrogen flushing has been used successfully for this purpose. A wide variety of reducing agents and compounds to sequester metals and inhibit chain reactions has been employed for stabilization, but relatively few are acceptable for parenteral products. Often, it is necessary to combine ingredients and adjust pH to maximize stability. Detailed kinetic studies have been reported for the oxidative decomposition of prednisolone.

The physical pharmacist has a difficult task with oxidative degradation. Initially, experiments must be designed that will encompass many variables. Preparing samples at several concentrations containing antioxidants plus sequestering agents at several pH levels and placing them in flint or amber containers with and without nitrogen is a common procedure. The subsequent evaluation of these limited data is critical. Light-sensitivity studies with several formulations of prochlorperazine resulted in the selection of a stable formula. In a study with idoxuridine it was shown that placing the aqueous solution in an amber container was sufficient to protect the product from oxidative degradation.

Drug Substance–Excipient Interaction—Drug substance–excipient studies are designed to determine a list of excipients that can be used routinely in the final dosage forms. Lactose, sucrose, calcium sulfate, dicalcium phosphate, starch and magnesium stearate are some of the substances routinely tested in combinations. Some basic observations with the drug substance and/or its salt form sometimes can dictate what excipients can be used. For example, one would not consider using sucrose or lactose if the drug substance being considered is a primary amine. This system has the potential for interaction to form a colored compound readily detected by a color change.

Various means have been used for detecting potential interactions and incompatibilities. Diffuse reflectance techniques have been used to detect interactions. This has been done by comparing the spectra obtained initially with those obtained after storage at exaggerated conditions. A shift in absorption has been interpreted as an interaction. Thin-layer chromatography also has been used. When excipients are present it is usually advisable to set a mixture of the excipients at the same conditions as the excipient-drug mixtures. This will give a comparison of the chromatograms of both systems. If any new degradation products are present, the source may be determined more easily.

Mixtures containing at least two levels of drug concentration with excipients are sealed in vials containing 5% water. These vials are stored under exaggerated conditions of light and heat for various time periods. The resultant samples are observed physically and analyzed by an appropriate technique to get a qualitative determination. At this point in the stability evaluation, which is a preliminary screening process, it is not necessary to know exactly how much has degraded. It is an all-or-none effect. The search is for the excipients that have no effect on the stability of the active ingredient.

When solution interactions are being investigated and no incompatibilities are evident, it is wise to recommend an *in vivo* experiment to evaluate availability. On occasion, interaction may occur in solution that is not detectable with routine procedures. For example, clindamycin was found to interact with cyclamates, which interfere with the absorption of the drug.

Other Changes—Optically active substances may lose their optical activity; eg, through racemization. If the enantiomorphic compounds possess different degrees of physiologic action, such changes may result in reduced therapeutic effects. Epinephrine has been shown to undergo racemization under various acidic and basic conditions. Although the potential for this to become evident during a preformulation evaluation is rare, one should always be aware of this possibility. Polymerization is also a remote possibility. Darkening of glucose solution is attributed to polymerization of the breakdown product, 5-(hydroxymethyl)furfural. Isomerization, which is the process involving the change of one structure into another having the same empirical formula but with different properties in one or more respects, also can occur; again, the occurrence is rare. Deamination and decarboxylation can occur sometimes. This type of change would be detected easily since the resultant degradation products would have completely different properties.

Permeability

A preformulation evaluation should include studies to assess the passage of drug molecules across biological membranes. These membranes act as lipid barriers to most drugs and permit the absorption of lipid-soluble substances by passive diffusion. Lipid-insoluble substances can cross the barrier only with considerable difficulty. The pH-partition theory explains the interrelationship of the dissociation constant, lipid solubility, pH at the absorption site and the absorption characteristics of drugs across membranes. The theory has evolved following a series of investigations in laboratory animals and man and is the basis of much of the current understanding of absorption of drugs.

Data obtained from basic physical-chemical studies described earlier may give the preformulation scientist an indication of possible absorption difficulties. Experimental techniques are available that can be used to give a more accurate assessment of absorption problems. An *in vitro* system that has been used extensively consists of an aqueous/organic solvent/aqueous system which has the advantage of being simple, allows for accurate pH control, membrane thickness and other variables. It can be described mathematically in precise terms. However, the interpretation and correlation of data are limited when applied to biologic systems.

Another *in vitro* procedure, the everted sac technique, is a simple and reproducible method for determining the absorption characteristics of drugs. Isolated segments of rat small intestines are everted and filled with a solution of the drug being evaluated, and the passage of drug through the membrane is determined. This technique has been used to measure the permeability of a number of drug substances.[14] It also can evaluate both passive and active transport of drugs. The fact that the preparation has been removed from the animal and its normal blood supply is a distinct disadvantage.

The *in situ* technique developed by Doluisio, *et al*,[15] for the study of membrane permeability appears to overcome the disadvantages of the everted sac technique. Since the intestine is not removed from its blood supply, the results would be expected to be similar to those obtained in intact animals. A disadvantage of the technique is that the procedure does not account for the loss of fluid from the solution by absorption in the intestine. Nonabsorbable markers, such as phenol red, can be added to the drug solution to solve this problem.

The techniques described can give the preformulation scientist an indication of possible absorption problems or suggest that little or no difficulty will be observed in the passage of a particular drug product through the biological membranes. This information, along with eventual studies in man, serves to establish possible *in vitro/in vivo* correlation for dissolution and bioavailability. These data are important in establishing quality-control specifications for the products which will ensure consistent biological performance from subsequent lots.

Proteins and Peptides

Proteins and peptides produced by the commercialization of biotechnology are presenting preformulation scientists with new challenges. In general, protein and peptide drugs are more expensive to produce, more potent and more difficult to analyze than nonprotein and nonpeptide drugs. They frequently are formulated as parenterals instead of oral dosage forms because they are unable to be absorbed from the GI tract, unstable in GI fluids or subject to rapid first-pass metabolism. Degradation of proteins and peptides occurs not only by covalent bond reaction but also by denaturation. The prediction of shelf-life by the Arrhenius equation is usually not applicable.

Degradation by reaction of the covalent bond can be characterized by the following major reactions: hydrolysis, transpeptidation, racemization, oxidation, diketopiperazine formation, disulfide exchange and photodecomposition. Hydrolysis can occur at the peptide linkage (R-NH-CO-R), but it is more stable than the ester linkage (R-O-CO-R)

unless cleavage is assisted by a neighboring group. Hence, peptides such as oxytocin and captopril are stable enough for liquid parenteral formulations. Transpeptidation occurs when amino acid residues cyclize back onto the peptide chain and the cyclic intermediate undergoes hydrolysis. Racemization can occur in acidic or alkaline medium, and if proline or glycine occur in the *N*-terminal position, diketopiperazine formation is facilitated. Cysteine, methionine and tryptophan are susceptible to oxidation, and since disulfide exchange is concentration-dependent, oligomers are formed frequently as a result of the creation of disulfide bonds between peptide chains. Photodecomposition of tryptophan residues may lead to discoloration and photoproducts of increased molecular weight.

Degradation via denaturation occurs when the conformational structure of a protein or peptide is altered. Potential factors that can denature a molecule include ionic strength, surface-active agents or processing conditions that subject the molecule to shear or adsorption. Identification of the preferred conformation, and mechanisms by which it can be altered, is critical in formulating the molecule as a stable drug. Hydrogen bonds act to stabilize conformational structure and the presence of water promotes hydrogen bonding. Hence, agents that disrupt the water-protein interaction such as salts and molecules with ionic side chains can promote conformational instability.

Several methods can be used to study denaturation of proteins. These include thermal analysis, determination of critical micelle concentration, determination of cloud-point, light scattering and fluorescence spectrometry. Thermal analysis with a scanning microcalorimeter is used to measure energies of transition in solution and is useful for determining the effect of stabilizing excipients on proteins in solution. Measurement of the critical micelle concentration also can be used as a tool to study the ability of an excipient to stabilize or disrupt the hydrophobic interactions which promote micellerization. Cloud-point measurements (the temperature, when cooled, at which a solution becomes cloudy) also have been suggested as a tool to study the effects of solvents or excipients on denaturation. Fluorescence spectrometry can be used to measure thermal denaturation by using a fluorescent probe whose fluorescence increases when a protein is denatured.

Proteins and peptides can be stabilized in many ways, usually employing empirical, rather than theoretical, procedures. For parenteral formulations, excipients are added to enhance stability. Serum albumin, itself a relatively stable protein, is used commonly as a stabilizer for peptides and proteins. It may inhibit surface adsorption and act as a cryoprotectant during lyophilization. Amino acids, such as glutamic or aspartic acid, may chelate metals such as zinc, which may cause aggregation; however, metal ions, such as calcium, are essential to the stability of certain amylases and proteases. Phospholipids and fatty acids also are potential stabilizers. Even though surfactants have a high denaturing effect, they also may inhibit the effects of other denaturants.

Proteins, as opposed to nonprotein drugs, may find a dilute aqueous medium unfavorable. Therefore, one should attempt to create an environment similar to the natural habitat of the specific protein. This environment would be rich in proteins and carbohydrates, low in oxygen and have a high degree of immobilized water. However, as methodologies for studying denaturation and degradation become more defined, the number of excipients needed to stabilize a formulation can be limited selectively.

Formulation Ingredients

Although preliminary screening of commonly used excipients with new-drug substances has become routine in preformulation studies, there are occasions when problems arise because of the interaction with additives such as preservatives, stabilizers, dyes and, possibly, flavors. A discussion of some problems that have risen is in order to make formulators aware that they should be concerned about the potential for interaction whenever another ingredient is added to a formulation.

Preservatives—Each time a liquid or semisolid pharmaceutical dosage form is prepared, it is necessary to include a preservative in the formulation. Such preservatives as sodium benzoate, sorbic acid and the methyl and propyl esters of *p*-hydroxybenzoic acid (parabens) have been used in these systems for many years. There have been reports that the parabens have been inactivated when used in the presence of various surface-active agents and vegetable gums. This loss of activity might be due to the formation of complexes between the preservative and the surfactant. A dialysis technique has been used to demonstrate an interaction between polysorbate 80 and the parabens. This observation becomes critical if the level of preservative added is borderline with respect to the preservative-activity threshold. The desired preservative effect may not be achieved unless an excess of the preservative is added to compensate for that which is complexed. It also has been shown that molecular complexes form when the parabens are mixed with polyethylene glycol, methylcellulose, polyvinylpyrrolidone or gelatin. The degree of binding was less than that observed with polysorbate 80. Sorbic acid also interacts with polysorbates but does not interact with polyethylene glycols. The quaternary ammonium compounds also are bound by polysorbate 80 to reduce their preservative activity. Benzyl alcohol also was shown to be adsorbed by certain types of rubber stoppers. Subsequent work has shown that butyl rubber does not interact with benzyl alcohol.

Antioxidants—During the preformulation evaluation of compounds that are sensitive to oxidation often it is commonplace to test several levels of antioxidant concentrations added to aqueous systems in order to determine the relative effectiveness of the antioxidants. Sodium bisulfite and ascorbic acid are two antioxidants that are used widely in pharmaceutical systems. Sodium bisulfite yields a colorless water-soluble salt when it is oxidized. It will add to double bonds, react with aldehydes and certain ketones and contributes in bisulfite cleavage reactions. Many of the reactions with bisulfite are irreversible, and the resulting sulfonic acids frequently are biologically inactive. Epinephrine has been shown to interact with bisulfite to form a bisulfite addition product. Other sympathomimetic drugs, principally the *ortho*- or *para*-hydroxybenzyl alcohol derivatives, also react with bisulfite in a similar manner. The *meta*-hydroxy alcohol does not react. Sometimes these interactions are reversible as in the case with the adrenocorticosteroid molecules.

Ascorbic acid, on the other hand, is less reactive. However, when mixed with compounds having a primary amine nucleus, there is the tendency for interaction to form a highly colored Schiff base. One must be aware of this possibility when selecting a suitable antioxidant.

Suspending Agents—Occasionally, it will be necessary to consider the use of a suspending agent to prepare some preliminary suspension preparations for stability evaluation prior to starting toxicity testing. The physical pharmacist should be aware of the potential for these additives to react with the drug substance being evaluated. Anionic water-soluble compounds, such as sodium carboxymethylcellulose, alginic acid, carrageenin and other hydrocolloids, although generally considered inert, frequently interact with drug compounds in solution. Carboxymethylcellulose and carrageenin form complexes, or possibly salts, with many medicinal agents including procaine, chlorpromazine, benadryl,

quinine, chlorpheniramine, neomycin and kanamycin. In some instances the formation of the complex imparted better stability to the system. When this problem is suspected, it is important to conduct appropriate tests to insure that an interaction does not take place in the system being evaluated.

Dyes—Although preformulation tests usually are conducted long before any consideration of coloring the intended dosage forms, they should not be overlooked. Dyes are chemical in nature and contain reactive sites capable of causing incompatibilities. Several studies have demonstrated that certified dyes do react with drug substances. Sugars, such as dextrose, lactose and sucrose, were found to increase the rate of fading of FD&C Blue #2. Insoluble complexes also were formed when quaternary ammonium compounds were formulated with FD&C Blue #1.

Summary

The preformulation evaluation of new-drug substances has become an integral part of the development process. A thorough understanding of the physical-chemical properties of the new-drug substance under study provides the development pharmacist with data that are essential in designing stable and efficacious dosage forms. Many of the problems discussed and the solutions offered in this chapter resulted from application of scientific training of present-day pharmaceutical scientists. Their diverse skills, creative aptitudes and initiative provide the pharmaceutical industry with the essential ingredients to develop drug products that help maintain the health-care process at its highest level of excellence.

References

1. Dittert LW, *et al: J Pharm Sci 57:* 1146, 1968.
2. Dittert LW, *et al: J Pharm Sci 57:* 1269, 1968.
3. Halebian H, McCrone W: *J Pharm Sci 58:* 911, 1969.
4. Ravin LJ, Higuchi T: *J APhA, Sci Ed 46:* 732, 1957.
5. Aguiar AJ, Zelmer JE: *J Pharm Sci 58:* 983, 1969.
6. Spiegel AJ, Noseworthy MM: *J Pharm Sci 52:* 917, 1963.
7. Guttman DE, *et al: J Pharm Sci 50:* 305, 1961.
8. Berge SM, *et al: J Pharm Sci 66:* 1, 1977.
9. Fell J, Newton J: *J Pharm Sci 59:* 688, 1970.
10. Hiestand H, Smith D: *Powder Tech 38:* 145, 1984.
11. Ravin LJ, *et al: J Pharm Sci 67:* 1523, 1978.
12. Ravin LJ, *et al: J Pharm Sci 53:* 1064, 1964.
13. Rattie ES, Guttman DE, Ravin LJ: *Arzneim-Forsch 28:* 944, 1978.
14. Kaplan SA, Colder S: *J Pharm Sci 61:* 1361, 1972.
15. Doluisio JT, Billups NF, Dittert LW, Sugita ET, Swintosky JV: *J Pharm Sci 58:* 1196, 1969.

Bibliography

Carstensen JT: *Pharmaceutics of Solids and Solid Dosage Forms*, Wiley, New York, 1977.
Greene DS: *Modern Pharmaceutics*, Marcel Dekker, New York, Chap 6, 1979.
Poole JW: *FMC Problem Solver and Reference Manual*, FMC Corp, Philadelphia, Section 5, 1982.
Fiese EF, Hagen TA: *The Theory and Practice of Industrial Pharmacy*, Lea & Febiger, Philadelphia, 1986, Chap 8.
Leuenberger H, Rohera BD, *Pharm Res 3:* 12, 1986.
Wang YJ, Hanson MA, *J Parenteral Sci Technol 42* (2S): S3, 1988.

CHAPTER 76

Bioavailability and Bioequivalency Testing

Anthony R DiSanto, PhD
Vice President, Drug Delivery Research and Development
The Upjohn Company
Kalamazoo, MI 49001

Pharmacy is a profession that requires the use of a number of scientific disciplines as well as the individual professional experience of its practitioners. Compounding of medications has become a small part of the pharmacist's practice, now replaced largely by his major role and responsibility for safeguarding drug-product quality through proper selection of multisource drug products. One need not become embroiled in the controversy of brand-name vs generic products, for this is not the issue. The problem is one of discriminate selection of a drug product available from different manufacturers—often of substitution of one product for another, whether it involves a brand-to-generic, generic-to-brand or generic-to-generic change.

For the pharmacist to accept such responsibility, he must be reasonably knowledgeable in biopharmaceutics, with particular emphasis on drug bioavailability and bioequivalence. Variable clinical response to the same dosage form of a drug product supplied by two or more drug manufacturers is well-recognized. In this chapter only bioavailability problems will be discussed. Chemical equivalence, lot-to-lot uniformity of physicochemical characteristics and stability equivalence are but a few of the other factors that are important, as they too can affect a patient's ultimate clinical response to a drug.

One must not be led to a feeling of overconfidence in the simplicity of product selection solely because the FDA promulgated bioavailability regulations. Even for the limited number of multisource drug products that require some type of bioequivalence testing, it should be recognized that the testing is only on one lot of the product. Similarly, where only *in vitro* assessment is required, data provided are limited to one to three lots. There is a misconception that once a product is marketed that the FDA continues to test each lot. This is not the case as very few drug products are followed up at the FDA laboratories. The question of *continued* assurance of bioequivalence and chemical equivalence must, therefore, be posed by the pharmacist. This is where the challenge lies, and the pharmacist has to call on both his technical training and experience to make appropriate drug-product selection decisions.

Bioavailability

In any discussion of bioavailability and bioequivalency testing, it is perhaps best to start with the basic concepts and factors that can affect the bioavailability of a drug and consider how these can affect bioequivalency and the clinical outcome of drug treatment. At the outset, the terms used in this chapter require careful definition since, as in any area, some terms have been used in many different contexts by different authors.

Bioavailability is an absolute term that indicates measurement of both the true rate and total amount (extent) of drug that reaches the general circulation from an administered dosage form.

Equivalence is more a relative term that compares one drug product with another or with a set of established standards. Equivalence may be defined in several ways:

1. *Chemical equivalence* indicates that two or more dosage forms contain the labeled quantities (plus or minus specified range limits) of the drug.
2. *Clinical equivalence* occurs when the same drug from two or more dosage forms gives identical *in vivo* effects as measured by a pharmacological response or by control of a symptom or disease.
3. *Therapeutic equivalence* implies that one structurally different chemical can yield the same clinical result as another chemical.
4. *Bioequivalence* indicates that a drug in two or more similar dosage forms reaches the general circulation at the same *relative* rate and the same *relative* extent, ie, that the plasma (blood or serum) level profiles of the drug obtained using the two dosage forms are, within reason, "superimposable."

Dosage Forms—In the dose titration of any patient the objective is, in conceptual terms, to attain and maintain a blood level which exceeds the minimum effective level required for response, but which does not exceed the minimum toxic (side-effect) level. This is shown graphically in Fig 76-1. There are three major absorption factors which can affect the general shape of this blood-level curve and thus drug response.

1. The dose of the drug administered, ie, the blood levels will rise and fall in proportion to the dose administered.
2. The same as the first but brought about by a different process, is the amount of drug absorbed from a given dosage form. The effect of having only one-half of the drug absorbed from a dosage form is equivalent to lowering the dose (Fig 76-2).
3. The rate of absorption of the drug. If absorption from the dosage form is more rapid than the rate of absorption which gave the profile in Fig 76-1, toxic (side-effect) levels can be exceeded. If absorption from the dosage form is sufficiently slow, minimum effective levels may never be attained (Fig 76-3).

A combination of these last two factors is also possible (Fig 76-4) and is probably the most likely result in real life.

In any of these instances, the time course and extent of clinical response to the drug has been altered.

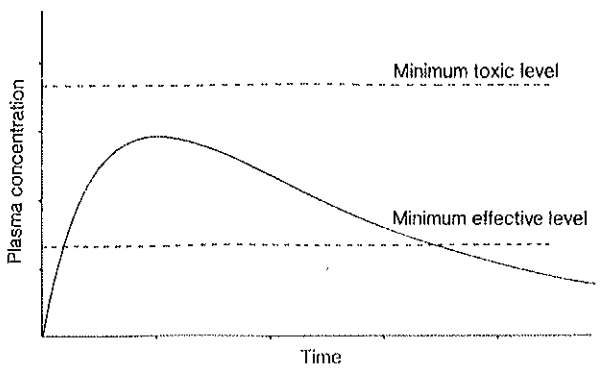


Fig 76-1. Typical plasma-level curve of a drug with effective and toxic (side-effect) levels defined.

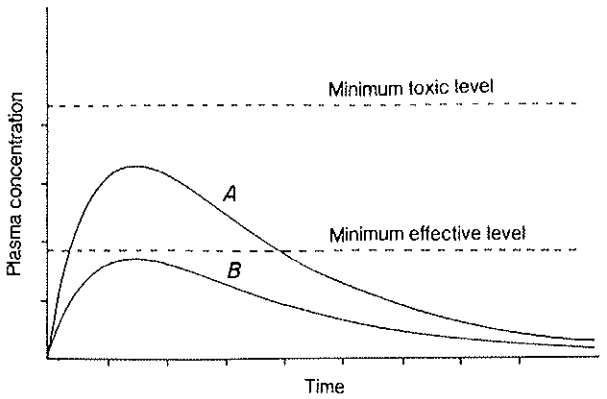


Fig 76-2. Effect of the extent of drug absorption from a dosage form on drug-plasma levels and efficacy. The extent of absorption from Dosage Form *B* is 50% of that from Dosage Form *A*.

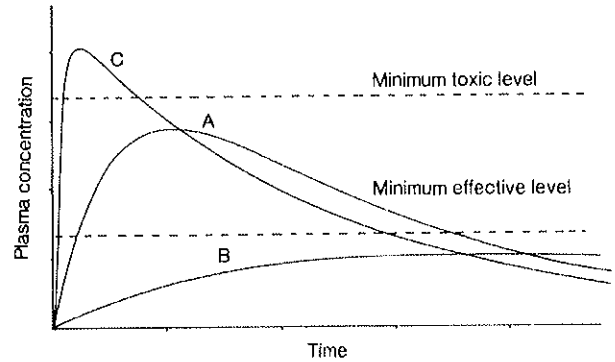


Fig 76-3. Effect of the rate of drug absorption from a dosage form on the plasma-level profile and efficacy. The rates of absorption from Dosage Forms *B* and *C* are 1/10 and 10 times those from Dosage Form *A*.

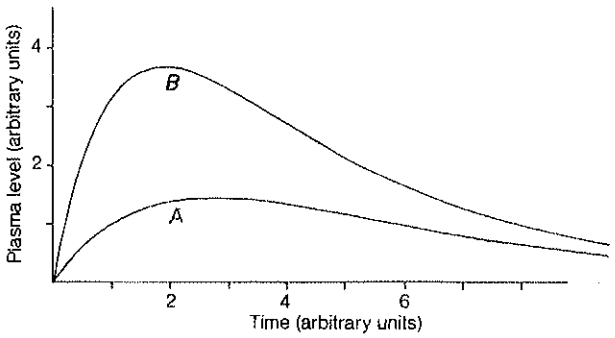


Fig 76-4. Computer simulation of the plasma-level curves for two dosage forms of the same drug assuming that the rate and extent of drug absorption for Dosage Form *A* were 50% and 50%, respectively, of those for Dosage Form *B*.

Both factors, extent and rate of drug absorption, can be affected by the dosage form in which the drug is contained. The effect may be intentional, as in sustained-release medication, or unintentional, as brought about by a change in the composition and/or method of manufacture of the dosage form.

It is important to remember that in most dosage forms the only ingredient regulated by law is the active drug. The choice of the other materials (adjuvants) used to prepare a satisfactory dosage form is up to the individual manufacturer. It is through these changes, in composition and manufacturing technique, that unintended changes in bioavailability and bioequivalency may occur. A description of the formulation of dosage forms and the factors which must be considered by the formulating pharmacist is given in Chapter 75.

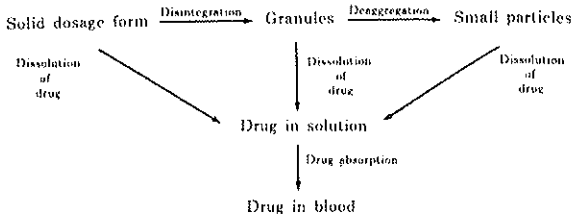


Fig 76-5. Sequence of events involved in the dissolution and absorption of a drug from a solid oral dosage form.

Dissolution Rate—For a drug to be absorbed, it must first go into solution. In Fig 76-5, the steps in the dissolution and absorption of a tablet or capsule dosage form are outlined. Similar profiles could be developed for any solid or semisolid dosage form, ie, oral suspensions, parenteral suspensions or suppositories. The theory and mechanics of drug-dissolution rate are described in detail in Chapter 31. Suffice it to say that the physical characteristics of the drug and the composition of the tablet (dosage form) can have an effect on the rates of disintegration, deaggregation and dissolution of the drug. As such, these can affect the rate of absorption and resultant blood levels of the drug.

Properties of the Drug—The physical characteristics of the drug which can alter bioavailability are discussed in Chapters 35 and 75 and consist of: the polymorphic crystal form, choice of the salt form, particle size, use of the hydrated or anhydrous form, wettability and solubility of the drug. Chapter 75 also discusses several other properties which can affect drug-product quality adversely. Many of these factors should be discovered during the chemical testing of the drug product prior to the sale of the dosage form and should not, therefore, affect, unknowingly, the bioavailability of the drug product.

Properties of the Dosage Form—The various components of the solid or semisolid dosage form, other than the active ingredient, are discussed in Chapter 89. Only an overview, for tablet dosage forms, will he given here. In addition to the active ingredient, a tablet product usually will contain:

Binders are used to provide a free-flowing powder from the mix of tablet ingredients so that the material will flow when used on a tablet machine. The binder also provides a cohesiveness to the tablet. Too little binder will give flow problems and tablets which do not maintain their integrity; too mueh may affect adversely the release (dissolution rate) of the drug from the tablet.

Fillers are used to give the powder bulk so that an acceptable-size tablet is produced. Most commercial tablets weigh from 100 to 500 mg so it is obvious that for many potent drugs the filler comprises a large portion of the tablet. The binding of drug to the filler may occur and affect bioavailability.

Disintegrants are used to cause the tablets to disintegrate when exposed to an aqueous environment. Too much will produce tablets which may disintegrate in the bottle due to atmospheric moisture; too little may be insufficient for disintegration to occur and may thus alter the rate and extent of release of the drug from the dosage form.

Lubricants are used to enhance the flow of the powder to the tablet machine and to prevent sticking of the tablet in the die of the tablet machine after the tablet is compressed. Lubricants are usually hydrophobic materials such as stearic acid, magnesium or calcium stearate. Too little lubricant will not permit satisfactory tablets to be made; too mueh may produce a tablet with a water-impervious hydrophobic coat, which can inhibit the disintegration of the tablet and dissolution of the drug.

The integrity of the manufacturer is not a true physical ingredient of the tablet, but it can have an effect on the clinical performance of the dosage form. Many of the problems which arise here are related to, and detectable hy, the physical and chemical quality controls the manufacturer applies to his product (see Chapter 82). For example, with low-dose potent drugs the determination that all the active ingredient is present, on the average, in the dosage form must be complemented by

the determination that each tablet contains the specified dose. It is quite possible with potent drugs that the assay of combined tablets (10 to 20) may be within compendial limits while the drug contents of individual tablets may far exceed these limits in both positive and negative directions. Such variations in dose, and thus bioavailability, are detectable and controllable by a chemical assay of the tablets. However, these assays and other determinations may not always be done by manufacturers with low integrity. This defect may be out of ignorance of the law or intentional disregard for it. The existence of laws and federal regulations does not mean that everyone, at any given point in time, is complying with such laws and regulations.

Bioequivalency Testing

The awareness of the potential for clinical differences between otherwise chemically equivalent drug products has been brought about by a multiplicity of factors which include, among others, better methods for clinical efficacy evaluation, development of techniques to measure microgram or nanogram quantities of drugs in biological fluids, improvements in the technology of dosage-form formulation and physical testing, awareness of a significant number of reported clinical inequivalencies in the literature, increased costs of classical clinical evaluation, the objective, quantitative nature of bioavailability tests and the increase in the number of chemically equivalent products on the market due to patent expirations on the wonder drugs of the 1950s and 1960s.

The increase in the number of similar products from multiple sources frequently has placed people involved in the delivery of health care in the position of having to select one from among several apparently equivalent products. As with any decision, the more pertinent the data available, the more comfortable one is in arriving at the final decision. The need to make these choices, in light of the potential for *in vivo* inequivalency among products, has increased the demand for quantitative data on the clinical equivalence of similar drug products. Bioequivalency testing represents one alternative solution to clinical testing for efficacy.

Requirements for bioequivalency data on drug products should not be applied indiscriminately. For example, with single-supplier drugs, for which clinical efficacy has been established, bioequivalency testing is moot. However, bioavailability data on three lots would be an excellent measure of reproduceable bioavailability. This assures the quality of the innovator and should serve as a guide for permissable variability in the multiscore product. In this context the *raison d'etre* for bioequivalency testing should not be forgotten, ie, it has been developed to substitute for the clinical evaluation of drug products. Bioequivalency data cannot be required if bioanalytical methodology is not available. However, in a number of cases pharmacodynamic data may provide a more sensitive, objective evaluation of a product's clinical equivalency than will clinical testing.

Pharmacokinetic evaluation of bioavailability data is not necessary to show bioequivalence of two drug products. Pharmacokinetics has its major utility in the prediction or projection of dosage regimens and/or in providing a better understanding of observed drug reactions or interactions which result from the accumulation of drug in some specific site, tissue or "compartment" of the body. The basis of all statements that two drug products are bioequivalent must be that the responses observed (blood, serum or plasma level, urinary excretion or pharmacologic response) for one drug product essentially are superimposable on the responses observed for the second drug product.

The phrase "essentially are superimposable" must be consistent with the clinical realities of the situation. The easy, but relatively rare, decisions in the evaluation of the bioequivalence of two drug products are those where the two products are exactly superimposable (definitely bioequivalent) and those where the two products differ in their bioequivalency parameters by 50% or more (definitely *bioinequivalent*). The demonstration of absolute differences of 10% or less in the bioavailability of two dosage forms is an assignment which frequently is not possible with today's analytical tools and clinical facilities. In the area of 10 to 20% or even 30% differences between two dosage forms in bioequivalency parameters, clinical judgment must be applied to evaluate the significance of these differences. The effect of a possible 10 to 30% change in dose on the patient's response must be considered carefully before one decides that an apparent or possible 20% difference in bioavailability is acceptable *or* unacceptable. It should be noted that the usual bioavailability difference allowed by the FDA is ±20%. There is no absolute reason why this value was picked.

Even with dosage forms whose bioavailabilities have been established (within 10 to 20%), there is a potential for undesirable, unexpected clinical response when changing the medication for a well-stabilized patient from one drug supplier to another.

It is important to realize that a 10 to 20% bioavailability difference observed in normal, healthy volunteers cannot be any less in a patient where factors affecting drug absorption already may be compromised. These relatively small bioavailability differences observed in healthy volunteers could be doubled or tripled depending on the disease, the state of the disease, the age of the patient, whether the patient is bedridden, has achlorhydria, has hypermotility or hypomobility, etc. Variables associated with the patient in general are unreconcilable and their individual cumulative effect on bioavailability is unknown. When one compounds this patient variability with a drug product that is less than optimally absorbed, the outcome cannot be predicted. The patient for whom the drug is prescribed is the critical factor not to be overlooked in product selection.

Evaluation of Bioequivalency Data

The following sections will highlight some of the tests that should be considered when evaluating the data from bioequivalency studies. The topics discussed will be directed specifically toward blood- or plasma-level evaluations. With minor modifications, the approaches outlined can be used for urinary excretion measurements or for suitable, quantitative pharmacological response measurements.

General Study Design—Bioavailability studies usually are conducted in normal, healthy adults under standardized conditions. Usually, single doses of the test and reference product will be evaluated. However, in selected cases, multiple-dose regimens must be used, eg, acid-labile drugs. The goal of the studies is to evaluate the performance of the dosage forms under standardized conditions. The assumption that any change in conditions or subject health will affect both dosage forms in a similar fashion is not valid and separate tests should be performed.

The protocol should define the acceptable age and weight range for the subjects to be used. It should define the clinical parameters which will be used to characterize a normal, healthy adult; eg, physical examination observations, clinical chemistry and hematological evaluations. The subjects should have been drug-free for at least 2 weeks prior to testing to eliminate possible drug-induced influences on liver enzyme systems. Normally, the subjects will fast overnight prior to dosing and will not eat until a standard meal is provided 2 to 4 hr postdosing. The dosage forms should be given to subjects in a randomized manner, using a suitable crossover design, so that possible daily variations are distributed equally between all dosage forms tested. The protocol should define sample-collection times and techniques to collect the biological fluid. The method of storage of the samples also should be defined.

Bioavailability Assessment and Data Evaluation—Several parameters are used to provide a general evaluation of the overall rate and extent of absorption of a drug. An analysis of all characteristics is required before one can implicate any one factor or parameter as indicating bioequivalence or a lack of bioequivalence.

The blood (or serum or plasma) concentration-time curve is the focal point of bioavailability assessment and is obtained when serial blood samples taken after drug administration are analyzed for drug concentration. The concentrations are plotted on graph paper on the ordinate (or y) axis and the times after drug administration that the samples were obtained on the abscissa (or x) axis.

A drug product is administered orally at time zero, and the blood drug concentration at this time clearly should be zero. As the product passes through the gastrointestinal system (stomach, intestine) it must go through the sequence of events depicted in Fig 76-5. As the drug is absorbed, increasing concentrations of the drug are observed in successive samples until the maximum concentration is achieved. This point of maximum concentration is called the peak of the concentration-time curve. If a simple one-compartment model describes the pharmacokinetics of the drug tested, the peak concentration represents approximately the point in time when absorption and elimination of the drug have equalized.

The section of the curve to the left of the peak represents the absorption phase (usually absorption and distribution), during which the rate of absorption exceeds the rate of elimination. The section of the curve to the right of the peak is called the elimination phase, during which the rate of elimination exceeds the rate of absorption. It should be understood that elimination begins as soon as the drug appears in the blood stream and continues until all of the drug has been eliminated. Elimination is classically the log-linear portion of the curve. Absorption continues too for some period of time into the elimination phase.

One must recognize that elimination of the drug includes all processes of elimination, urinary excretion as well as metabolism, of the drug by various tissues and organs. The "efficiency" of metabolism and urinary excretion will determine the shape of the elimination phase of the curve.

Bioavailability studies are performed in healthy, adult volunteers under rigid conditions of fasting and activity because the objective is to obtain *quantitative* information on the influence of pharmaceutical formulation variables on the drug-product's absorption. Drug blood-level profiles, therefore, allow quantification of the *rate* and *extent* of drug absorption and are critical in establishing the *efficiency* of the drug product in delivering the drug to the systemic circulation.

Arguments that bioavailability testing should be done in a "disease-state population" are not tenable if the object of the study is to assess drug formulations. If, on the other hand, the purpose is to determine the effect of "disease" on the efficiency of absorption from the drug product(s), then one must use the disease-state population. The reasoning is obvious. In order to assure that any differences observed in the drug blood-level profiles are attributable to formulation factors, one must hold all other variables constant, ie, food, activity, state of disease, etc.

One need not be limited to drug blood-level profiles, but in a similar manner many obtain cumulative urinary drug amount-time profiles. Drug *concentration* is determined in the urine at specified time intervals and the *amount* excreted per interval determined by multiplying the concentration by the volume of urine obtained in that interval. The amounts per interval then are cumulated and ultimately the maximum amount excreted in the urine is obtained. This value is analogous to the area under the blood concentration-time curve. A typical cumulative urinary drug amount-time profile for several nitrofurantoin products is presented in Fig 76-6.

In assessing the bioequivalency of drug products one must quantitate the *rate* and *extent* of absorption. The factors of the rate and extent of absorption can be determined by evaluating three parameters of a blood level concentration-time profile. Three parameters describing a blood level curve are considered important in evaluating the bioequivalency of two or more formulations of the same drug; these are the peak height concentration, the time of the peak concentration and the area under the blood (serum or plasma) concentration-time curve.

Peak Height Concentration—The height of the peak of the blood level-time curve obviously represents the highest drug concentration achieved after oral administration. It is reported as an amount per volume measurement, eg, micrograms/mL or units/mL or grams/100 mL, etc. The importance of this parameter is illustrated in Fig 76-7 where the blood concentration-time curves of two different formulations of a drug are represented. A line has been drawn across the curve at 4 μg/mL. Suppose the drug is an analgesic and 4 μg/mL is the minimum effective concentration (*MEC*) of the drug in blood. If, then, the blood concentration curves in Fig 76-7 represent the blood levels obtained after administration of equal doses of two formulations of the drug and it is known that analgesia would not be produced unless the minimum effective concentration was achieved or exceeded, it becomes clear that Formulation *A* should produce pain relief while Formulation *B*, even though it seemed well-absorbed, would not produce the desired pharmacological effect and would be ineffective in producing analgesia.

On the other hand, if the two curves represent blood con-

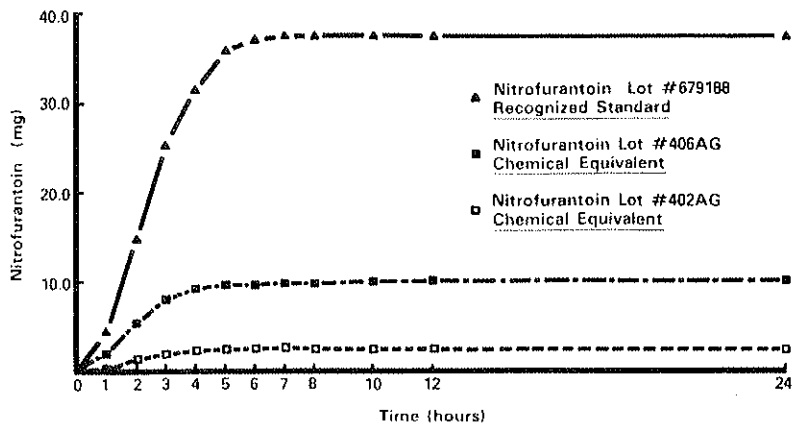


Fig 76-6. Average cumulative amounts of nitrofurantoin excreted from three lots of two commercially available products after a single oral dose of 100 mg of nitrofurantoin.

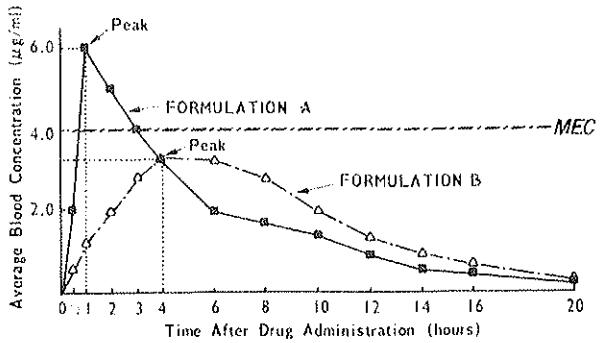


Fig 76-7. Blood concentration-time curves obtained for two different formulations of the same drug demonstrating relationship of the profiles to the minimum effective concentration (*MEC*).

centrations following equal doses of two different formulations of the same cardiac glycoside, and 4 μg/mL now represents the minimum toxic concentration (*MTC*) and 2 μg/mL represents the *MEC* (Fig 76-8), Formulation *A*, although effective, may also be toxic, while Formulation *B* produces concentrations well above the *MEC* but never achieves toxic levels.

Time of Peak Concentration—The second parameter of importance is the measurement of the length of time necessary to achieve the maximum concentration after drug administration. This time is called the time of peak blood concentration. In Fig 76-7, for Formulation *A* the time necessary to achieve peak blood concentration is 1 hr; for Formulation *B* it is 4 hr. This parameter is related closely to the rate of absorption of the drug from a formulation and may be used as a simple measure of rate of absorption.

To illustrate its importance, suppose the two curves in Fig 76-8 now represent two formulations of an analgesic and that in this case the minimum effective concentration is 2 μg/mL. Formulation *A* will achieve the *MEC* in 30 min; Formulation *B* does not achieve that concentration until 2 hr. Obviously, Formulation *A* would then produce analgesia much more rapidly than Formulation *B* and would probably be preferable as an analgesic agent. On the other hand, if one were more interested in the duration of the analgesic effect than on the time of onset, Formulation *B* would present more sustained activity, maintaining serum concentrations above the *MEC* for a longer time (8 hr) than Formulation *A* (5½ hr).

Area Under the Concentration-Time Curve—The third, and sometimes the most important parameter for evaluation, is the area under the serum, blood or plasma concentration-time curve (AUC). This area is reported in amount/volume × time (eg, μg/mL × hours or grams/100 mL × hours, etc) and can be considered representative of the amount of drug absorbed following administration of a single dose of the drug.

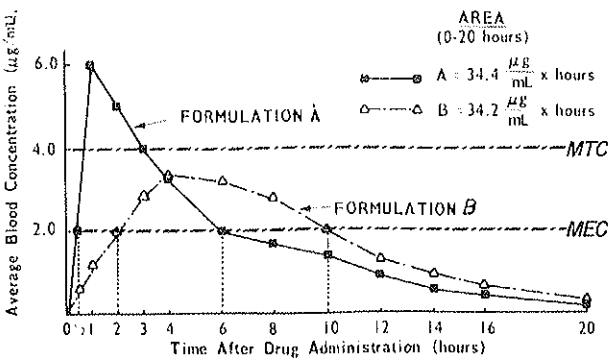


Fig 76-8. Blood concentration-time curves obtained for two different formulations of same the drug demonstrating relationship of the profiles to the minimum toxic concentration (*MTC*) and the minimum effective concentration (*MEC*).

Returning to Fig 76-8, the curves, although much different in shape, have approximately the same areas (A = 34.4 μg/mL × hours; B = 34.2 μg/mL × hours) and both formulations can be considered to deliver the same amount of drug to the systemic circulation. Thus, one can see that AUC does not represent the only criterion on which bioequivalency can be judged. All the results, as a composite, must be used in reaching a decision as to bioequivalency; no one parameter serves this purpose.

Statistical Sense and Nonsense—When statistical evaluations are employed in bioequivalency testing one must be careful not to assume, from a statement that "no statistically significant differences were detected," that two drug products are, therefore, bioequivalent. The basis of most tests for statistically significant differences is that the two products are assumed to be the same until proven otherwise. Therefore, if the data presented are highly variable (large standard deviation, ie, wide range of values), it would be possible to show that there was no statistically significant difference between an AUC of 100 units (%) versus an AUC of 40 units (%). In this case the statistical test does not indicate that the AUCs are truly similar; it simply means that the data were too variable from patient to patient for the statistics to be able to detect a 60-unit (%) difference in areas, even if it existed.

There are two types of errors associated with any statistical test. These are:

1. **Alpha (α) Error**—This is the error with which most people are familiar and is the error associated with the statement, "The data have been analyzed statistically." α error is the probability (defined by the p value) by saying the two treatments are different when in fact they are the same. It should be noted that while highly significant p values reduce the alpha error, they provide no indication of the possibility that the two treatments being called the same when in fact they are different.
2. **Beta (β) Error**—This is the error associated with the possibility of calling two treatments the same when in fact they are different. As the maximum percent difference between means which can be detected with an α error of $p \leq 0.05$ is reduced, the β error also is reduced. This increase in statistical sensitivity (reduced α and β error) is obtained by reducing the variability of the data. Variability usually is reduced by increasing the number of data points (subjects) in a bioavailability study. It is implicit that the analytical methodology is specific, sensitive and precise.

The objective of statistical testing for bioavailability evaluation should be to minimize both the α and β error. Since both errors are related mathematically to the variability of the data collected, the solution is relatively simple. Sufficient data should be gathered so that the general statistical test (α error test) would detect, if it existed, a predetermined percent difference (20% for example) between the two dosage forms. If, for example, the two treatments are found statistically not to be ($p \leq 0.05$) different significantly, the results indicate that there is only 1 chance in 20 that the treatments are claimed to be different when in fact they are the same.

If there were 18 subjects in the above example and a 20% difference would have been significantly different statistically, there would be a β error of 4 chances in 20 that a 25% difference between means was not detected. That is, that treatments which differed by more than 25% were claimed to be the same when in fact they were different. The level of statistical sensitivity which one feels is adequate (20% as a rule of thumb) must be reevaluated for each drug product tested based on the clinical performance of the drug.

Statistical analysis also can go to the other extreme. For example, tests might show that an AUC of 100 units (100%) was statistically significantly different from an AUC of 90

units (90%). If the clinical impression of the drug being evaluated was that a 20% difference in dose (plasma levels) would not be clinically significant, in this example it must be concluded that the statistical test is too sensitive and the difference observed, even if real, is not significant clinically. Therefore, the drug products are bioequivalent in spite of the statistical findings.

Statistics should be used, in bioavailability testing, as a tool to determine if sufficient subjects have been included to minimize the effect of patient-to-patient variability in the data analysis. The results of statistical testing should not be used *as the decision* but to help *make the decision*. One must apply some statistical sense in order to avoid statistical nonsense.

A Common Pitfall: Cross-Study Comparisons—Perhaps the single most-common error made in interpreting bioavailability data is that of *cross-study comparison*. This occurs when the blood concentration-time curve of a drug product in one study is compared with the blood concentration-time curve of that drug product in another study. There are three reasons why such cross-study comparisons are dangerous and can lead to false conclusions. The following examples used to illustrate the three points are taken from actual bioavailability data.

Different Subject Population—In Fig 76-9, a research lot of potassium phenoxymethyl penicillin was compared with the appropriate reference standard for that product. The research-lot drug was found to be bioequivalent, with average peak-serum concentrations differing by 8% and the area differing by only 9%. In another study conducted with a full-manufacture lot of the test product, the same lot of the reference standard potassium phenoxymethyl penicillin was used. The results of this study are shown in Fig 76-10. Again, the two products were found to be bioequivalent as the peak and area parameters differed by less than 5%. In these two studies, identical test conditions were used and the same analytical procedure and laboratory was employed. However, if one compares the serum levels for the reference standard lot found in Fig 76-9, with the levels for the same lot of tablets in the study in Fig 76-10, sizable differences in blood levels are found as shown in Fig 76-11.

The average peak serum levels for this lot of tablets were found to be 8.5 units/mL and 12.5 units/mL in the two respective studies, a difference of approximately 31%. Likewise, the average AUC was found to differ by approximately 21%. Such differences are the sole result of cross-study

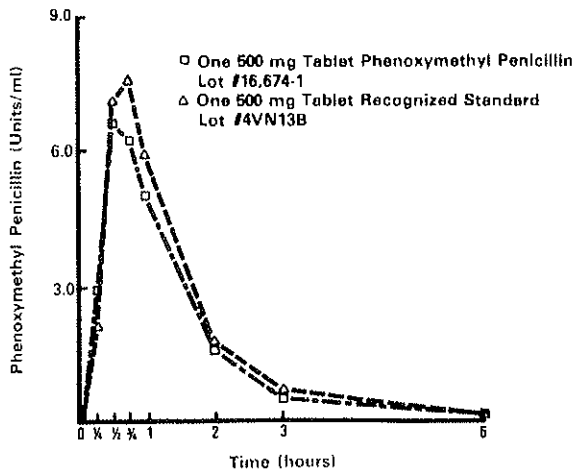


Fig 76-9. Average serum concentration of phenoxymethyl penicillin following oral administration of 500 mg given as one tablet of Recognized Standard (△), or of Test Product, Research Lot (□).

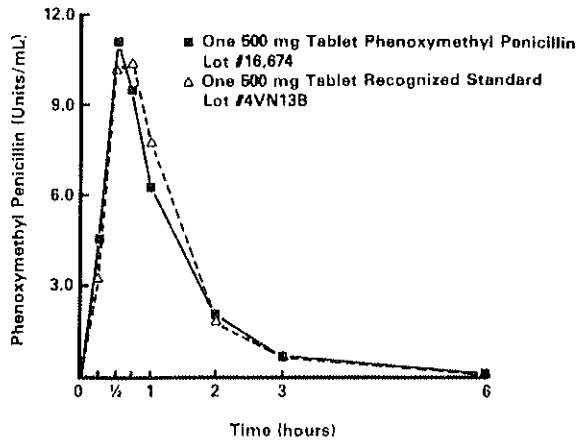


Fig 76-10. Average serum concentration of phenoxymethyl penicillin following oral administration of 500 mg given as one tablet of Recognized Standard (△), or of Test Product, Full Mfg Lot (■).

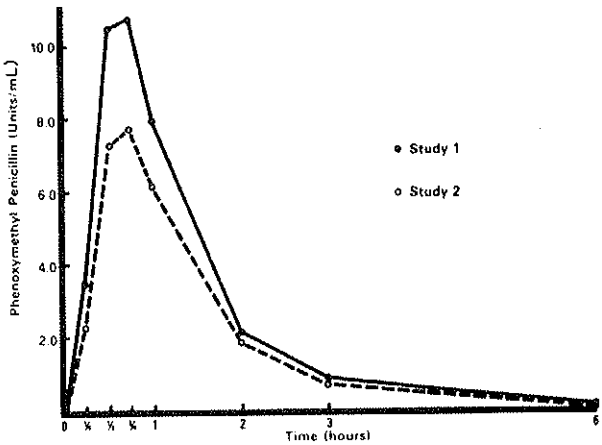


Fig 76-11. Average serum concentration of phenoxymethyl penicillin following a single oral 500-mg dose of Recognized Standard, in two different subject populations.

comparisons and are not due to differences in actual bioavailability.

The same lot of reference standard tablets was used in both studies. Hence, the difference must be due to the experimental variables which occur normally from study to study. The major difference between the two studies was the subject population involved. In the first study, healthy, adult, male, prison volunteers were used, whereas in the second study, there were 17 females and 7 males in a hospital clinic, also described as normal, healthy volunteers. An appreciable difference in sex distribution was obvious when comparing these studies. Adjustments for body weight and surface area alone did not correct for the apparent discrepancies in peak concentration or blood level AUC. It is difficult to determine the exact factors which caused the observed differences. This example should serve as a note of caution in comparing absolute bioavailability values of peak concentration and area under the curve from different studies.

Different Study Conditions—Parameters such as the food or fluid intake of the subject before, during and after drug administration can have dramatic effects on the absorption of certain drugs. Fig 76-12 shows the results of a three-way crossover test where the subjects were fasted 12 hr overnight and 2 hr after drug administration of an uncoated

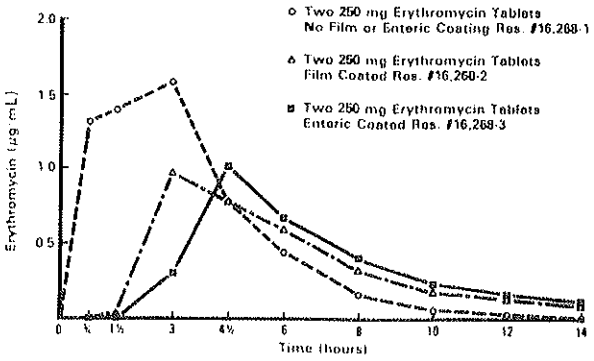


Fig 76-12. Average serum erythromycin concentration administered in 500-mg doses as three different tablet dosage forms. The results were obtained from 21 healthy adult subjects following an overnight fast of 12 hr before and 2 hr after drug administration.

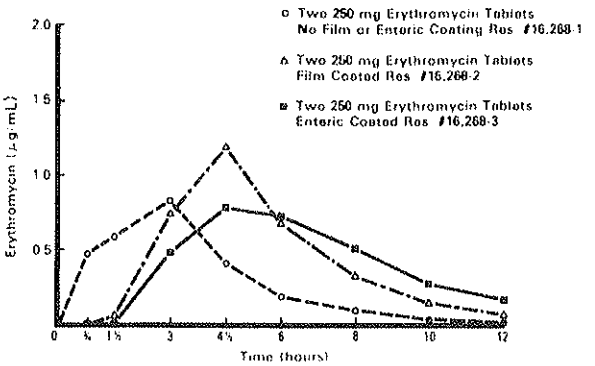


Fig 76-13. Average serum erythromycin concentration administered in 500-mg doses as three different tablet dosage forms. The results were obtained from 12 healthy adult subjects with only a 2-hr fast before drug administration.

tablet, a film-coated tablet or an enteric-coated tablet of erythromycin.

The results of this study suggest that the unprotected tablet is superior to both the film-coated and enteric-coated tablets in terms of blood-level performance. These results

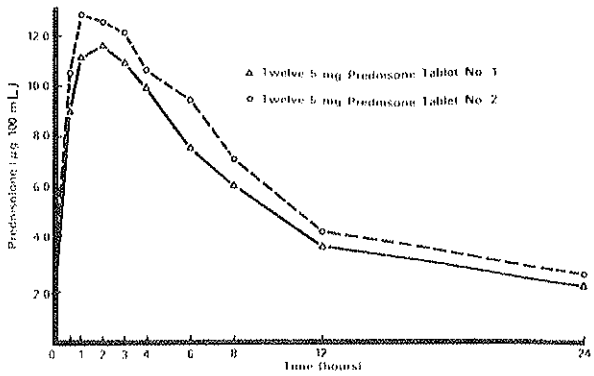


Fig 76-15. Average plasma prednisolone levels following 60 mg of prednisone administered to 24 normal adults as a single oral dose of twelve 5-mg prednisone tablets from two different manufacturers. Plasma levels were determined by a competitive protein-binding assay.

also suggest that neither film-coating nor enteric-coating is necessary for optimal blood-level performance. Figure 76-13 shows results with the same tablets when the study conditions were changed to only a 2-hr preadministration fast with a 2-hr postadministration fast. In this case, the blood levels of the uncoated tablet were depressed markedly while the film-coated and enteric-coated tablets showed relatively little difference in blood levels.

From this second study, it might be concluded that film-coating appears to impart the same degree of acid stability as an enteric coating. This might be acceptable if only one dose of the antibiotic was required. However, Fig 76-14 shows the results of a multiple-dose study in which the enteric-coated tablet and the film-coated tablet were administered 4 times a day, immediately after meals. The results show that the film coating does not impart the degree of acid stability as does the enteric coating when the tablets are administered immediately after food in a typical clinical situation.

Different Assay Methodology—Depending on the drug under study, there may be more than one assay method available. For example, some steroids can be assayed by a radioimmunoassay, competitive protein-binding, gas-liquid

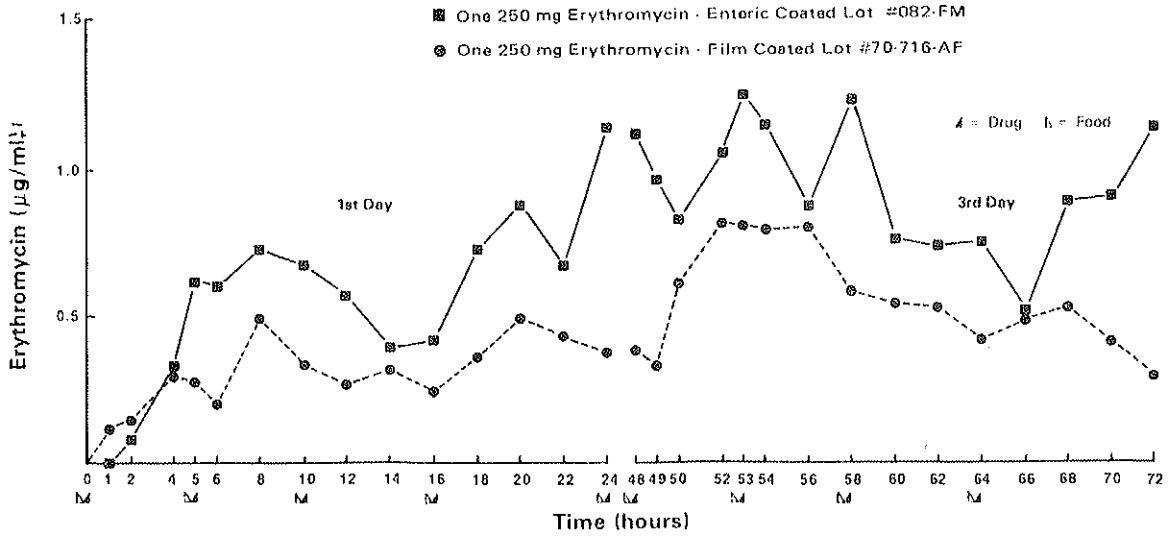


Fig 76-14. Average serum erythromycin concentration-time profiles administered in two different tablet dosage forms. The results were obtained from 24 healthy adult subjects following administration of 250 mg 4 times a day, with meals and at bedtime.

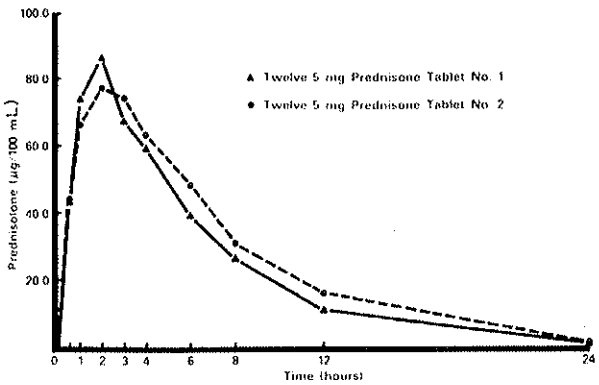


Fig 76-16. Average plasma prednisolone levels following 60 mg of prednisone administered to 24 normal adults as a single oral dose of 12 5-mg prednisone tablets from two different manufacturers. Plasma levels were determined by a radioimmunoassay procedure.

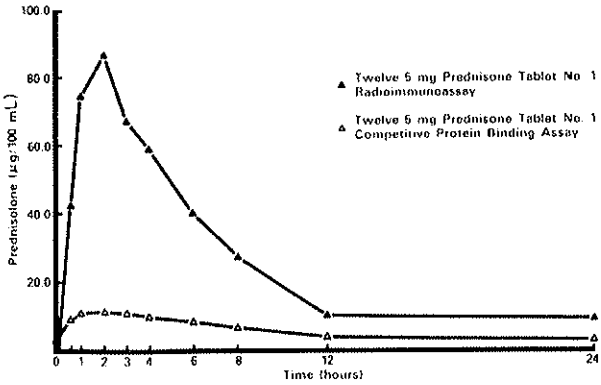


Fig 76-17. Average plasma prednisolone profiles administered as a single 60-mg dose to 24 normal adults. Plasma levels were determined by both a competitive protein binding assay and a radioimmunoassay.

chromatograph or, indirectly, by a 17-hydroxycorticosteroid assay.

Figures 76-15 and 76-16 show the results of a comparison of prednisone tablets using a competitive protein-binding method and a radioimmunoassay, respectively. The serum concentration-time curves resulting from each method lead to the same conclusion, that the products are bioequivalent. However, Fig 76-17 shows a comparison of the absolute values obtained by the two assay methods with the same product.

Obviously, the wrong conclusion would have been reached if one product had been assayed by one method and the other product by the other method and the results had been compared. Even in cases where only one assay method is employed, there are numerous modifications with respect to technique among laboratories which could make direct comparisons hazardous.

The backbone of any bioavailability study involving plasma (or urine) levels of drug, in addition to good study design and subject controls, is the analytical methodology used to determine the levels of a drug. In most cases one *probably* can assume that the precision and reliability of the method employed in a given study have been established to a sufficient degree to make the results of the study internally consistent. As demonstrated, major problems arise when, without careful evaluation of the analytical methodology employed, one attempts to compare the data of studies from different laboratories. Even with similar analytical methodology performed by the same laboratory, it would be unreasonable to expect agreement, using the same dosage form, of closer than 20 to 25% for plasma levels, AUCs, etc, from one study to the next.

Under the *best* conditions, cross-study comparisons are relatively insensitive, and at worst they can be misleading. Cross-study comparisons certainly cannot be used to make decisions or estimations of differences in drug products with the generally acceptable sensitivity of difference detection of 20% or less.

With insufficient data on the correlation of plasma levels with clinical response, it is difficult to decide if it is the peak plasma level or the total body load of a drug that is important. Changes in the rate of absorption require changes in the dose given (body load) for maintenance of similar peak plasma levels. Decisions as to which is more important, body load or peak level, are made with difficulty and tend to reduce the objective quantitation sought in bioavailability testing.

Bibliography

Chodos DJ, DiSanto AR: *Basics of Bioavailability*, The Upjohn Co, Kalamazoo, 1973.
Dittert LW, DiSanto AR: *J APhA NS13*:(8): 000 1973.
DiSanto AR *et al*: *Int J Clin Pharmacol 13*: 220–227, 1976.

CHAPTER 79

Tonicity, Osmoticity, Osmololity and Osmolarity

Frederick P Siegel, PhD

Professor of Pharmaceutics
College of Pharmacy, University of Illinois
Chicago, IL 60612

It generally is accepted that osmotic effects have a major place in the maintenance of homeostasis (the state of equilibrium in the living body with respect to various functions and to the chemical composition of the fluids and tissues, eg, temperature, heart rate, blood pressure, water content or blood sugar). To a great extent these effects occur within or between cells and tissues where they cannot be measured. One of the most troublesome problems in clinical medicine is the maintenance of adequate body fluids and proper balance between extracellular and intracellular fluid volumes in seriously ill patients. It should be kept in mind, however, that fluid and electrolyte abnormalities are not diseases, but are the manifestations of disease.

The physiological mechanisms which control water intake and output appear to respond primarily to serum osmoticity. Renal regulation of output is influenced by variation in rate of release of pituitary antidiuretic hormone (ADH) and other factors in response to changes in serum osmoticity. Osmotic changes also serve as a stimulus to moderate thirst. This mechanism is sufficiently sensitive to limit variations in osmoticity in the normal individual to less than about 1%. Body fluid continually oscillates within this narrow range. An increase of plasma osmoticity of 1% will stimulate ADH release, result in reduction of urine flow and, at the same time, stimulate thirst that results in increased water intake. Both the increased renal reabsorption of water (without solute) stimulated by circulating ADH and the increased water intake tend to lower serum osmoticity.

The transfer of water through the cell membrane occurs so rapidly that any lack of osmotic equilibrium between the two fluid compartments in any given tissue usually is corrected within a few seconds and, at most, within a minute or so. However, this rapid transfer of water does not mean that complete equilibration occurs between the extracellular and intracellular compartments throughout the entire body within this same short period of time. The reason is that fluid usually enters the body through the gut and then must be transported by the circulatory system to all tissues before complete equilibration can occur. In the normal person it may require 30 to 60 min to achieve reasonably good equilibration throughout the body after drinking water. Osmoticity is the property that largely determines the physiologic acceptability of a variety of solutions used for therapeutic and nutritional purposes.

Pharmaceutical and therapeutic consideration of osmotic effects has been, to a great extent, directed toward the side effects of ophthalmic and parenteral medicinals due to abnormal osmoticity, and to either formulating to avoid the side effects or finding methods of administration to minimize them. More recently this consideration has been extended to total (central) parenteral nutrition, to enteral hyperalimentation ("tube" feeding) and to concentrated-fluid infant formulas.[1] Also, in recent years, the importance of osmometry of serum and urine in the diagnosis of many pathological conditions has been recognized.

There are a number of examples of the direct therapeutic effect of osmotic action, such as the intravenous use of mannitol as a diuretic which is filtered at the glomeruli and thus increases the osmotic pressure of tubular urine. Water must then be reabsorbed against a higher osmotic gradient than otherwise, so reabsorption is slower and diuresis is observed. The same fundamental principle applies to the intravenous administration of 30% urea used to affect intracranial pressure in the control of cerebral edema. Peritoneal dialysis fluids tend to be somewhat hyperosmotic to withdraw water and nitrogenous metabolites. Two to five percent sodium chloride solutions or dispersions in an oleaginous base (Muro, *Bausch & Lomb*) and a 40% glucose ointment are used topically for corneal edema. Ophthalgan (*Ayerst*) is ophthalmic glycerin employed for its osmotic effect to clear edematous cornea to facilitate an ophthalmoscopic or gonioscopic examination. Glycerin solutions in 50 to 75% concentrations [Glyrol (*IO Lab*), Osmoglyn (*Alcon*)] and isosorbide solution [Ismotic (*Alcon*)] are oral osmotic agents for reducing intraocular pressure. The osmotic principle also applies to plasma extenders such as polyvinylpyrrolidone and to saline laxatives such as magnesium sulfate, magnesium citrate solution, magnesium hydroxide (via gastric neutralization), sodium sulfate, sodium phosphate and sodium biphosphate oral solution and enema (*Fleet*).

An interesting osmotic laxative which is a nonelectrolyte is a lactulose solution. Lactulose is a nonabsorbable disaccharide which is colon-specific, wherein colonic bacteria degrade some of the disaccharide to lactic and other simple organic acids. These, *in toto*, lead to an osmotic effect and laxation. An extension of this therapy is illustrated by Cephulic (*Merrell-Dow*) solution, which uses the acidification of the colon via lactulose degradation to serve as a trap for ammonia migrating from the blood to the colon. The conversion of ammonia of blood to the ammonium ion in the colon ultimately is coupled with the osmotic effect and laxation, thus expelling undesirable levels of blood ammonia. This product is employed to prevent and treat frontal systemic encephalopathy.

Osmotic laxation is known with the oral or rectal use of glycerin and sorbitol. Epsom salt has been used in baths and compresses to reduce edema associated with sprains. A relatively new approach is the indirect application of the osmotic effect in therapy via osmotic pump drug delivery systems.[2]

If a solution is placed in contact with a membrane that is permeable to molecules of the solvent, but not to molecules of the solute, the movement of solvent through the membrane is called osmosis. Such a membrane is often called *semipermeable*. As the several types of membranes of the body vary in their permeability, it is well to note that they are *selectively* permeable. Most normal living-cell membranes maintain various solute concentration gradients. A selectively permeable membrane may be defined either as one that does not permit free, unhampered diffusion of all

the solutes present, or as one that maintains at least one solute concentration gradient across itself. Osmosis, then, is the diffusion of water through a membrane that maintains at least one solute concentration gradient across itself.

Assume a Solution A on one side of the membrane, and a Solution B of the same solute but of a higher concentration on the other side; the solvent will tend to pass into the more concentrated solution until equilibrium has been established. The pressure required to prevent this movement is the osmotic pressure. It is defined as the excess pressure, or pressure greater than that above the pure solvent, which must be applied to Solution B to prevent passage of solvent through a perfect semipermeable membrane from A to B. The concentration of a solution with respect to effect on osmotic pressure is related to the number of particles (unionized molecules, ions, macromolecules, aggregates) of solute(s) in solution and thus is affected by the degree of ionization or aggregation of the solute. See Chapter 16 for review of colligative properties of solutions.

Body fluids, including blood and lacrimal fluid, normally have an osmotic pressure which often is described as corresponding to that of a 0.9% solution of sodium chloride. The body also attempts to keep the osmotic pressure of the contents of the gastrointestinal tract at about this level, but there the normal range is much wider than that of most body fluids. The 0.9% sodium chloride solution is said to be *isoosmotic* with physiological fluids. The term *isotonic*, meaning equal tone, is in medical usage commonly used interchangeably with isoosmotic. However, terms such as isotonic and tonicity should be used *only* with reference to a physiologic fluid. Isoosmotic actually is a physical term which compares the osmotic pressure (or another colligative property, such as freezing-point depression) of two liquids, neither of which may be a physiological fluid, or which may be a physiological fluid only under certain circumstances. For example, a solution of boric acid that is isoosmotic with both blood and lacrimal fluid is isotonic only with the lacrimal fluid. This solution causes hemolysis of red blood cells because molecules of boric acid pass freely through the erythrocyte membrane regardless of concentration. Thus, isotonicity infers a sense of physiologic compatibility where isoosmoticity need not. As another example, a "chemically defined elemental diet" or enteral nutritional fluid can be isoosmotic with the contents of the gastrointestinal tract, but would not be considered a physiological fluid, or suitable for parenteral use.

A solution is isotonic with a living cell if there is no net gain or loss of water by the cell, or other change in the cell when it is in contact with that solution. Physiologic solutions with an osmotic pressure lower than that of body fluids, or of 0.9% sodium chloride solution, are referred to commonly as being *hypotonic*. Physiological solutions having a greater osmotic pressure are termed *hypertonic*.

Such qualitative terms are of limited value, and it has become necessary to state osmotic properties in quantitative terms. To do so, a term must be used that will represent all the particles which may be present in a given system. The term used is *osmol*. An osmol is defined as the weight, in grams, of a solute, existing in a solution as molecules (and/or ions, macromolecules, aggregates, etc), which is osmotically equivalent to a mole of an ideally behaving nonelectrolyte. Thus, the osmol-weight of a nonelectrolyte, in a dilute solution, generally is equal to its gram-molecular-weight. A milliosmol, abbreviated mOsm, is the weight stated in milligrams.

If one extrapolates this concept of relating an osmol and a mole of a nonelectrolyte as being equivalent, then one also may define an osmol in the following ways. It is the amount of solute which will provide one Avogadro's number (6.02×10^{23}) of particles in solution and it is the amount of solute which, on dissolution in 1 kg of water, will result in an osmotic pressure increase of 22.4 atmospheres. This is derived from the gas equation, PV = nRT, assuming ideal conditions and standard temperature of 0°. This is equivalent to an increase of 17,000 torr or 19,300 torr at 37°. One mOsmol is one-thousandth of an osmol. For example, 1 mole of anhydrous dextrose is equal to 180 g. One Osmol of this nonelectrolyte is also 180 g. One mOsmol would be 180 mg. Thus 180 mg of this solute dissolved in 1 kg of water will produce an increase in osmotic pressure of 19.3 torr at body temperature.

For a solution of an electrolyte such as sodium chloride, one molecule of sodium chloride represents one sodium and one chloride ion. Hence, one mole will represent 2 osmols of sodium chloride theoretically. Accordingly, 1 osmol NaCl = 58.5 g/2 or 29.25 g. This quantity represents the sum total of 6.02×10^{23} ions as the total number of particles. Ideal solutions infer very dilute solutions or infinite dilution. However, as the concentration is increased, other factors enter. With strong electrolytes, interionic attraction causes a decrease in their effect on colligative properties. In addition, and in opposition, for all solutes, including nonelectrolytes, solvation and possibly other factors operate to intensify their colligative effect. Therefore, it is very difficult and often impossible to predict accurately the osmoticity of a solution. It may be possible to do so for a dilute solution of a single, pure and well-characterized solute, but not for most parenteral and enteral medicinal and/or nutritional fluids; experimental determination likely is required.

Osmolality and Osmolarity

It is necessary to use several additional terms to define expressions of concentration in reflecting the osmoticity of solutions. The terms include osmolality, the expression of osmolal concentration and osmolarity, the expression of osmolar concentration.

Osmolality—A solution has an osmolal concentration of one when it contains 1 osmol of solute/kg of water. A solution has an osmolality of n when it contains n osmols/kg of water. Osmolal solutions, like their counterpart molal solutions, reflect a weight to weight relationship between the solute and the solvent. All solutions with the same molal concentrations, irrespective of solute, contain the same mole fraction (f_m) of solute. In water

$$f_m = \frac{\text{moles solute}}{\text{moles solute} + \text{moles solvent}}$$

thus, for a one molal solution

$$f_m = \frac{1 \text{ mole solute}}{1 \text{ mole solute} + 55.5 \text{ moles water per kg}} = \frac{1}{56.5}$$

Since an osmol of any nonelectrolyte is equivalent to 1 mole of that compound, then a 1 osmolal solution is synonymous to a 1 molal solution for a typical nonelectrolyte.

With a typical electrolyte like sodium chloride, 1 osmol is approximately 0.5 mole of sodium chloride. Thus, it follows that a 1 osmolal solution of sodium chloride essentially is equivalent to a 0.5 molal solution. Recall that a 1 osmolal solution of dextrose or sodium chloride each will contain the same particle concentration. In the dextrose solution there will be 6.02×10^{23} molecules/kg of water and in the sodium chloride solution one will have 6.02×10^{23} total ions/kg of water, one-half of which are Na^+ ions and the other half Cl^- ions. The mole fraction, in terms of total particles, will be the same and, hence, the same osmotic pressure.

As in molal solutions, osmolal solutions usually are employed where quantitative precision is required, as in the

measurement of physical and chemical properties of solutions (ie, colligative properties). The advantage of the *w/w* relationship is that the concentration of the system is not influenced by temperature.

Osmolarity—The relationship observed between molality and osmolality is shared similarly between molarity and osmolarity. A solution has an osmolar concentration of 1 when it contains 1 osmol of solute/L of solution. Likewise, a solution has an osmolarity of *n* when it contains *n* osmols/L of solution. Osmolar solutions, unlike osmolal solution, reflect a weight in volume relationship between the solute and final solution. A one molar and 1 osmolar solution would be synonymous for nonelectrolytes. For sodium chloride a 1 osmolar solution would contain 1 osmol of sodium chloride per liter which approximates a 0.5 molar solution. The advantage of employing osmolar concentrations over osmolal concentrations is the ability to relate a specific number of osmols or milliosmols to a volume, such as a liter or mL. Thus, the osmolar concept is simpler and more practical. The osmolal concept does not allow for this convenience because of the *w/w* relationship. Also, additional data such as the density usually are not available. Volumes of solution, rather than weights of solution, are more practical in the delivery of liquid dosage forms.

Many health professionals do not have a clear understanding of the difference between osmolality and osmolarity. In fact, the terms have been used interchangeably. This is due partly to the circumstance that, until recently, most of the systems involved were body fluids, in which the difference between the numerical values of the two concentration expressions is small and similar in magnitude, to the error involved in their determination. The problem partly may center around the interpretation by some to view one kilogram of water in the osmolal concept as being equivalent to 1 L, and, more importantly, the interpretation that to make up to volume of 1 L, as in osmolarity, is essentially the same as the weight of solute plus 1 liter (a distortion of the osmolal concept). The primary difference resides in the error introduced which revolves around the volume of water occupied by the solute. A 1 osmolar solution of a solute always will be more concentrated than a 1 osmolal solution. With dilute solutions the difference may be acceptably small. Nine grams of sodium chloride/L of aqueous solution is approximately equivalent to 9 g in 996.5 mL of water. This represents an error of under 1%, when comparing the osmoticity of 0.9% *w/v* solution to a solution of 9 g plus 1 kg of water. Using dextrose in a parallel comparison, errors range from approximately 3.5% in osmoticity with 50 g dextrose/L versus 50 g plus 1 kg of water to a difference of about 25% in osmoticity with 250 g dextrose/L versus 250 g plus 1 kg of water. The confusion appears to be without cause for concern at this time. However, one should be alerted to the sizeable errors which may occur with concentrated solutions or fluids, such as those employed in total parenteral nutrition, enteral hyperalimentation and oral nutritional fluids for infants.

Reference has been made to the terms hypertonic and hypotonic. Analogous terms are hyperosmotic and hypoosmotic. The significance of hyper- and hypo-osmoticity for medicinal and nutritional fluids will be discussed later. The values which correspond to those terms for serum may be visualized approximately from the following example. *Assuming* normal serum osmolality to be *285 mOsmol/kg*, as serum osmolality increases due to water *deficit*, the following signs and symptoms usually are found to accumulate progressively at approximately these values: 294 to 298—thirst (if the patient is alert and communicative); 299 to 313—dry mucous membranes; 314 to 329—weakness, doughy skin; above 330—disorientation, postural hypotension, severe weakness, fainting, CNS changes, stupor and coma. As serum osmolality decreases due to water *excess* the following may occur: 275 to 261—headache; 262 to 251—drowsiness, weakness; 250 to 233—disorientation, cramps; below 233—seizures, stupor and coma.

As indicated previously, the mechanisms of the body actively combat such major changes by limiting the variation in osmolality for normal individuals to less than about 1% (approximately in the range 282 to 286 mOsmol/kg, based on the above assumption).

The value given for normal serum osmolality above was described as an assumption because of the variety of values found in the literature. Serum osmolality often is stated loosely to be about 300 mOsmol/L. Apart from that, and more specifically, two references state it as 280 to 295 mOsmol/L; other references give it as 275 to 300 mOsmol/L, 290 mOsmol/L, 306 mOsmol/L, and 275 to 295 mOsmol/kg. There is a strong tendency to call it *osmolality* but to state it as *mOsmol/L* (*not as mOsmol/kg*). In the light of these varying values, one may ask about the reproducibility of the experimental measurements. It has been stated that most osmometers are accurate to 5 mOsmol/L. With that type of reproducibility, the above variations perhaps may be expected. The difference between a liter and kilogram probably is insignificant for serum and urine. It is difficult to measure kilograms of water in a solution, and easy to express body fluid quantities in liters. Perhaps no harm has been done to date by this practice for body fluids. However, loose terminology here may lead to loose terminology when dealing with the rather concentrated fluids used at times in parenteral and enteral nutrition.

Reference has been made to confusion in the use of the terms osmolality and osmolarity, a distinction of special importance for nutritional fluids. Awareness of high concentrations of infant-formula should give warning as to possible risks. Unfortunately, the osmoticity of infant formulas, tube feedings and total parenteral nutrition solutions has not been described adequately either in textbooks or in the literature,[3] and the labels of many commercial nutritional fluids do not, in any way, state their osmoticity. Only recently have enteral fluids been characterized in terms of osmoticity. Some product lines now are accenting isoosmotic enteral nutritional supplements. Often, when the term osmolarity is used, one cannot discern whether this simply is incorrect terminology, or if osmolarity actually has been calculated from osmolality.

Another current practice which can cause confusion, is the use of the terms *normal* and/or *physiological* for isotonic sodium chloride solution (0.9%). The solution surely is isoosmotic. However, as to being physiological, the concentration of ions are each of 154 mEq/L while serum contains about 140 mEq of sodium and about 103 mEq of chloride.

The range of mOsmol values found for serum raises the question as to what really is meant by the terms hypotonic and hypertonic for medicinal and nutritional fluids. One can find the statement that fluids with an osmolality of 50 mOsmol or more above normal are hypertonic and, if 50 mOsmol or more below normal, are hypotonic. One also can find the statement that peripheral infusions should not have an osmolarity exceeding 700 to 800 mOsmol/L.[4] Examples of osmol concentrations of solutions used in peripheral infusions are: D5W—252 mOsmol/L; D10W—505 mOsmol/L; Lactated Ringer's 5% Dextrose—525 mOsmol/L. When a fluid is hypertonic, undesirable effects often can be decreased by using relatively slow rates of infusion, and/or relatively short periods of infusion. D25W—4.25% Amino Acids is a representative example of a highly osmotic hyperalimentation solution. It has been stated that when osmolal loading is needed, a maximum safe tolerance for a normally hydrated subject would be an approximate increase of 25 mOsmol/kg of water over 4 hr.[5]

Computation of Osmolarity

Several methods are used to obtain numerical values of osmolarity. The osmolar concentration, sometimes referred to as the "theoretical osmolarity", is calculated from the wt/vol concentration using one of the following equations:

For a nonelectrolyte

$$\frac{\text{g/L}}{\text{mol wt}} \times 1000 = \text{mOsmol/L} \qquad (1)$$

For a strong electrolyte

$$\frac{\text{g/L}}{\text{mol wt}} \times \frac{\text{number of ions formed}}{} \times 1000 = \text{mOsmol/L} \qquad (2)$$

For individual ions, if desired

$$\frac{\text{g of ion/L}}{\text{ionic wt}} \times 1000 = \text{mOsmol (of ion)/L} \qquad (3)$$

These are simple calculations, however, they omit consideration of factors such as solvation and interionic forces. By this method of calculation 0.9% sodium chloride has an osmolar concentration of 308 mOsmol/L.

Two other methods compute osmolarity from values of osmolality. The determination of osmolality will be discussed later. One method has a strong theoretical basis of physical-chemical principles[6] using values of the partial molal volume(s) of the solute(s). A 0.9% sodium chloride solution, found experimentally to have an osmolality of 286 mOsmol/kg, was calculated to have an osmolarity of 280 mOsmol/L, rather different from the value of 308 mOsmol/L calculated as above. The method, using partial molal volumes, is relatively rigorous, but many systems appear to be too complex and/or too poorly defined to be dealt with by this method.

The other method is based on the following relationship:[6,7] actual osmolarity = measured osmolality × (density − g solute/mL). This expression can be written

$$\text{mOsmol/L solution} = \text{mOsmol/1000 g water} \times \text{g water/mL solution}$$

The experimental value for the osmolality of 0.9% sodium chloride solution was 292.7 mOsmol/kg; the value computed for osmolarity was 291.4 mOsmol/L. This method does not have as firm a theoretical basis as the preceding method but it has the advantage that it uses easily obtained values of density of the solution and of its solute content. Apparently, it can be used with all systems. For example, the osmolality of a nutritional product was determined by the freezing point depression method to be 625 mOsmol/kg;[7] its osmolarity was calculated as 625 × 0.839 = 524 mOsmol/L.

The USP requires that labels of pharmacopeial solutions which provide intravenous replenishment of fluid, nutrient(s), or electrolyte(s), as well as of the osmotic diuretic Mannitol Injection, state the osmolar concentration, in milliosmols/L, except that where the contents are less than 100 mL, or where the label states the article is not for direct injection but is to be diluted before use, the label alternatively may state the total osmolar concentration in milliosmols/mL. This is a reasonable request from several standpoints, and intravenous fluids are being labeled in accordance with this stipulation, as shown in the next section.

An example of the use of the first method described above is the computation of the approximate osmolar concentration ("theoretical osmolarity") of a Lactated Ringer's 5% Dextrose Solution (*Abbott*), which is labeled to contain, per L, dextrose (hydrous) 50 g, sodium chloride 6 g, potassium chloride 300 mg, calcium chloride 200 mg and sodium lactate 3.1 g. Also stated is that the total osmolar concentration of the solution is approximately 524 mOsmol per L, in part contributed by 130 mEq of Na^+, 109 mEq of Cl^-, 4 mEq of K^+, 3 mEq of Ca^{2+} and 28 mEq of lactate ion.

The derivation of the osmolar concentrations from the stated composition of the solution may be verified by calculations using Eq 1 above for the nonelectrolyte dextrose, and Eq 2 for the electrolytes.

Dextrose

$$\frac{50\text{ g} \times 1000}{198.17} = 252.3\text{ mOsmol/L}$$

Sodium Chloride

$$\frac{6\text{ g} \times 2 \times 1000}{58.44} = 205.33\text{ mOsmol/L} \begin{cases} (102.66\text{ mOsmol } Na^+) \\ (102.66\text{ mOsmol } Cl^-) \end{cases}$$

Potassium Chloride

$$\frac{0.3\text{ g} \times 2 \times 1000}{74.55} = 8.04\text{ mOsmol/L} \begin{cases} (4.02\text{ mOsmol } K^+) \\ (4.02\text{ mOsmol } Cl^-) \end{cases}$$

Calcium Chloride

$$\frac{0.2\text{ g} \times 3 \times 1000}{110.99} = 5.4\text{ mOsmol/L} \begin{cases} (1.8\text{ mOsmol } Ca^{2+}) \\ (3.6\text{ mOsmol } Cl^-) \end{cases}$$

Sodium Lactate

$$\frac{3.1\text{ g} \times 2 \times 1000}{112.06} = 55.32\text{ mOsmol/L} \begin{cases} (27.66\text{ mOsmol } Na^+) \\ (27.66\text{ mOsmol lactate}) \end{cases}$$

The total osmolar concentration of the five solutes in the solution is 526.4, in good agreement with the labeled total osmolar concentration of approximately 524 mOsmol/L.

The mOsmol of sodium in 1 L of the solution is the sum of the mOsmol of the ion from sodium chloride and sodium lactate, ie, 102.66 + 27.66 = 130.32 mOsmol. Chloride ions come from the sodium chloride, potassium chloride and calcium chloride, the total osmolar concentration being 102.66 + 4.02 + 3.6 = 110.3 mOsmol. The mOsmol values of potassium, calcium and lactate are calculated to be 4.02, 1.8 and 27.66, respectively. Thus, with the possible exception of calcium, there is close agreement with the labeled mEq content of each of these ions.

The osmolarity of a mixture of complex composition, such as an enteral hyperalimentation fluid, probably cannot be calculated with any acceptable degree of certainty and, therefore, the *osmolality* of such preparations probably should be determined experimentally.

The approximate osmolarity of mixtures of two solutions can be computed from the following relationship (the method is known as *alligation medial*)

$$\text{osm}_{\text{final}} = \frac{\text{osm}_a \times V_a}{V_{\text{final}}} + \frac{\text{osm}_b \times V_b}{V_{\text{final}}}$$

where

V_a = volume of component a
V_b = volume of component b
V_{final} = volume of final solution
osm_a = osmolarity of component a
osm_b = osmolarity of component b
$\text{osm}_{\text{final}}$ = osmolarity of final solution

For example, to calculate the osmolarity of a mixture of 500 mL of a solution of osmolarity 850 and 500 mL of a solution of osmolarity 252:

$$\text{osm}_{\text{final}} = \frac{850 \times 500}{1000} + \frac{252 \times 500}{1000}$$

$$= 425\text{ mOsmol/L} + 126\text{ mOsmol/L} = 551\text{ mOsmol/L}$$

This example illustrates the ease of calculating the osmoticity, by use of osmolarity, when solutions are mixed. Such a calculation would be much less valid if osmolality values were used. From the previous example one can see how to calculate the approximate effect if an additional solute is added.

Undesirable Effects of Abnormal Osmoticity

Ophthalmic Medication—It generally has been accepted that ophthalmic preparations intended for instillation

into the cul-de-sac of the eye should, if possible, be approximately isotonic to avoid irritation (see Chapter 86). It also has been stated that the abnormal tonicity of contact lens solutions can cause the lens to adhere to the eye and/or cause burning or dryness and photophobia.

Parenteral Medication—Osmoticity is of great importance in parenteral injections, its effects depending on the degree of deviation from tonicity, the concentration, the location of the injection, the volume injected, the speed of the injection, the rapidity of dilution and diffusion, etc. When formulating parenterals, solutions otherwise hypotonic usually have their tonicity adjusted by the addition of dextrose or sodium chloride. Hypertonic parenteral drug solutions cannot be adjusted. Hypotonic and hypertonic solutions usually are administered slowly in small volumes, or into a large vein such as the subclavian, where dilution and distribution occur rapidly. Solutions that differ from the serum in tonicity generally are stated to cause tissue irritation, pain on injection and electrolyte shifts, the effect depending on the degree of deviation from tonicity.

Excessive infusion of *hypo*tonic fluids may cause swelling of red blood cells, hemolysis and water invasion of the body's cells in general. When this is beyond the body's tolerance for water, water intoxication results, with convulsions and edema, such as pulmonary edema.

Excessive infusion of *iso*tonic fluids can cause an increase in extracellular fluid volume, which can result in circulatory overload.

Excessive infusion of *hyper*tonic fluids leads to a wide variety of complications. For example, the sequence of events when the body is presented with a large intravenous load of hypertonic fluid, rich in dextrose, is as follows: hyperglycemia, glycosuria and intracellular dehydration, osmotic diuresis, loss of water and electrolytes, dehydration and coma.

One cause of osmotic diuresis is the infusion of dextrose at a rate faster than the ability of the patient to metabolize it (as greater than perhaps 400 to 500 mg/kg/hr for an adult on total parenteral nutrition). A heavy load of nonmetabolizable dextrose increases the osmoticity of blood and acts as a diuretic; the increased solute load requires more fluid for excretion, 10 to 20 mL of water being required to excrete each gram of dextrose. Solutions, such as those for total parenteral nutrition, should be administered by means of a metered constant-infusion apparatus over a lengthy period (usually more than 24 hr) to avoid sudden hyperosmotic dextrose loads. Such solutions may cause osmotic diuresis; if this occurs, water balance is likely to become negative because of the increased urinary volume, and electrolyte depletion may occur because of excretion of sodium and potassium secondary to the osmotic diuresis. If such diuresis is marked, body weight falls abruptly and signs of dehydration appear. Urine should be monitored for signs of osmotic diuresis, such as glycosuria and increased urine volume.

If the intravenous injection rate of hypertonic solution is too rapid, there may be catastrophic effects on the circulatory and respiratory systems. Blood pressure may fall to dangerous levels, cardiac irregularities or arrest may ensue, respiration may become shallow and irregular and there may be heart failure and pulmonary edema. Probably the precipitating factor is a bolus of concentrated solute suddenly reaching the myocardium and the chemoreceptors in the aortic arch and carotid sinus.[5]

Abrupt changes in serum osmoticity can lead to cerebral hemorrhage. It has been shown experimentally that rapid infusions of therapeutic doses of hypertonic saline with osmotic loads produce a sudden rise in cerebrospinal fluid (CSF) pressure and venous pressure (VP) followed by a precipitous fall in CSF pressure. This particularly may be conducive to intracranial hemorrhage, as the rapid infusion produces an increase in plasma volume and venous pressure at the same time the CSF pressure is falling. During the CSF pressure rise, there is a drop in hemoglobin and hematocrit, reflecting a marked increase in blood volume.

Hyperosmotic medications, such as sodium bicarbonate (osmolarity of 1563 at 1 mEq/mL), which are administered intravenously, should be diluted prior to use and should be injected slowly to allow dilution by the circulating blood. Rapid "push" injections may cause a significant increase in blood osmoticity.[6]

As to other possibilities, there may be crenation of red blood cells and general cellular dehydration. Hypertonic dextrose or saline, etc, infused through a peripheral vein with small blood volume may traumatize the vein and cause thrombophlebitis. Infiltration can cause trauma and necrosis of tissues. Safety, therefore, demands that all intravenous injections, especially highly osmotic solutions, be performed slowly, usually being given preferably over a period not less than that required for a complete circulation of the blood, eg, *1 min*. The exact danger point varies with the state of the patient, the concentration of the solution, the nature of the solute and the rate of administration.

Hyperosmotic solutions also should not be discontinued suddenly. In dogs, marked increase in levels of intracranial pressure occur when hyperglycemia produced by dextrose infusions is reversed suddenly by stopping the infusion and administering saline. It also has been shown that the CSF pressure in humans rises during treatment of diabetic ketoacidosis in association with a fall in the plasma concentration of dextrose and a fall in plasma osmolality. These observations may be explained by the different rates of decline in dextrose content of the brain and of plasma. The concentration of dextrose in the brain may fall more slowly than in the plasma, causing a shift of fluid from the extracellular fluid space to the intracellular compartment of the CNS, resulting in increased intracranial pressure.

Osmometry and the Clinical Laboratory

Osmometry is a fairly recent innovation in the clinical laboratory; an article in 1971 had the title: "Osmometry: A New Bedside Laboratory Aid for the Management of Surgical Patients." Serum and urine osmometry may assist in the diagnosis of certain fluid and electrolyte problems. However, osmometry values have little meaning unless the clinical situation is known. Osmometry is used in renal dialysis as a check on the electrolyte composition of the fluid. In the clinical laboratory, as stated above, the term "osmolality" is used generally, but usually is reported as mOsmol/L. It may seem unnecessary to mention that osmolality depends not only on the number of solute particles, but also on the quantity of water in which they are dissolved. However, it may help one to understand the statement that the normal range of urine osmolality is 50 to 1400 mOsmol/L, and for a random specimen is 500 to 800 mOsmol/L.

Serum Osmoticity

Sodium is by far the principal solute involved in serum osmoticity. Therefore, abnormal serum osmoticity is most likely to be associated with conditions that cause abnormal sodium concentration and/or abnormal water volume.

Thus, hyperosmotic serum is likely to be caused by an increase in serum sodium and/or loss of water. It may be associated with diabetes insipidus, hypercalcemia, diuresis during severe hyperglycemia or with early recovery from renal shutdown. Alcohol ingestion is said to be the most common cause of the hyperosmotic state and of coexisting coma and the hyperosmotic state. An example of hyperos-

moticity is a comatose diabetic with a serum osmoticity of 365 mOsmol/L.

In a somewhat analogous fashion, hypoosmotic serum is likely to be due to decrease in serum sodium and/or excess of water. It may be associated with the postoperative state (especially with excessive water replacement therapy), treatment with diuretic drugs and low-salt diet (as with patients with heart failure, cirrhosis, etc), adrenal disease (eg, Addison's disease, adrenogenital syndrome) or SIADH (syndrome of inappropriate ADH secretion). There are many diseases which cause ADH to be released inappropriately (ie, in spite of the fact that serum osmoticity and volume may have been normal initially). These include oat-cell carcinoma of the lung, bronchogenic carcinoma, congestive heart failure, inflammatory pulmonary lesions, porphyria, severe hypothyroidism or cerebral disease (such as tumor, trauma, infection, vascular abnormalities, etc). It also may be found with some patients with excessive diuretic use. Serum and urine osmoticity are measured when SIADH is suspected. In SIADH there is hypoosmoticity of the blood in association with a relative hyperosmoticity of urine. The usual cause is a malfunction of the normal osmotic response of osmoreceptors, an excess of exogenous vasopressin, or a production of a vasopressin-like hormone that is not under the regular control of serum osmoticity. The diagnosis is made by simultaneous measurement of urine and serum osmolality. The serum osmolality will be lower than normal and much lower than the urine osmolality, indicating inappropriate secretion of a concentrated urine in the presence of a dilute serum.

Cardiac, renal and hepatic disease characteristically reduce the sodium/osmolality ratio, this being partially attributed to the effects of increased blood sugar, urea or unknown metabolic products. Patients in shock may develop disproportionately elevated measured osmolality compared to calculated osmolality, which points toward the presence of circulating metabolic products.

There are several approximate methods for estimating serum osmolality from clinical laboratory values for sodium ion, etc. They may be of considerable value in an emergency situation.

1. Serum osmolality may be estimated from the formula

$$\text{mOsmol} \simeq (1.86 \times \text{sodium}) + \frac{\text{blood sugar}}{18} + \frac{\text{BUN}}{2.8} + 5$$

(Na in mEq/L, blood sugar and BUN in mg/100 mL)

2. A quick approximation is

$$\text{mOsmol} = 2\,\text{Na} + \frac{\text{BS}}{20} + \frac{\text{BUN}}{3}$$

3. The osmolality is usually, *but not always*, very close to two times the sodium reading plus 10.

Urine Osmoticity

The two main functions of the kidney are glomerular filtration and tubular reabsorption. Clinically, tubular function is measured best by tests that determine the ability of the tubules to concentrate and dilute the urine. Tests of urinary dilution are not as sensitive in the detection of disease, as are tests of urinary concentration. As concentration of urine occurs in the renal medulla (interstitial fluids, loops of Henle, capillaries of the medulla and collecting tubules), the disease processes that disturb the function or structure of the medulla produce early impairment of the concentrating power of the kidney. Such diseases include acute tubular necrosis, obstructive uropathy, pyelonephritis, papillary necrosis, medullary cysts, hypokalemic and hypercalcemic nephropathy and sickle-cell disease.

Measurement of urine osmolality is an accurate test for the diluting and concentrating ability of the kidneys. In the absence of ADH, the daily urinary output is likely to be 6 to 8 L, or more. The normal urine osmolality depends on the clinical setting; normally, with maximum ADH stimulation, it can be as much as 1200 mOsmol/kg, and with maximum ADH suppression as little as 50 mOsmol/kg. Simultaneous determination of serum and urine osmolality often is valuable in assessing the distal tubular response to circulating ADH. For example, if the patient's serum is hyperosmolal, or in the upper limits of normal ranges, and the patient's urine osmolality measured at the same time is much lower, a decreased responsiveness of the distal tubules to circulating ADH is suggested.

Measurement of urine osmolality during water restriction is an accurate, sensitive test of decreased renal function. For example, under the conditions of one test, normal osmolality would be greater than 800 mOsmol/kg. With severe impairment the value would be less than 400 mOsmol/kg. Knowledge of urine osmolality may point to a problem even though other tests are normal (eg, the Fishberg concentration test, BUN, PSP excretion, creatinine clearance or IV pyelogram). Knowledge of its value may be useful especially in diabetes mellitus, essential hypertension and silent pyelonephritis. The urine/serum osmolality ratio should be calculated and should be equal to or greater than 3.

Osmoticity and Enteral Hyperalimentation

Some aspects of nutrition are discussed briefly here because of the potential major side effects due to abnormal osmoticity of nutritional fluids, and because there exists increasing dialogue on nutrition among pharmacists, dietitians, nurses and physicians. An example is the professional organization, ASPEN (The American Society for Parenteral and Enteral Nutrition), with membership open to all of the above health practitioners. It is desirable, therefore, that pharmacists be able to discuss these matters with these other health professionals in terms of nutrition as well as medicine.

Osmoticity has been of special importance in the intravenous infusion of large volumes of highly concentrated nutritional solutions. Their hyperosmoticity has been a major factor in the requirement that they be injected centrally into a large volume of rapidly moving blood, instead of using peripheral infusion. Use of such solutions and knowledge of their value seems to have led, more recently, to the use of rather similar formulations administered, not parenterally, but by instillation into some part of the gastrointestinal tract, usually, but not necessarily, by gavage. Of course, gavage feeding is not new. This method has given excellent total nutrition, for a period of time, to many patients. It has furnished an important part of their nutrition to others. It obviously avoids some of the problems associated with injections. Many of the reports on this topic refer to the use of a "Chemically Defined Elemental Diet." These are special nutritionally complete formulations that contain protein in so-called "elemental" or "predigested" form (protein hydrolysates or synthetic amino acids), and carbohydrate and fat in simple, easily digestible forms. These diets are necessarily relatively high in osmoticity because their smaller molecules result in more particles per gram than in normal foods. An example is a fluid consisting of: L-amino acids, dextrose oligosaccharides, vitamins (including fat-soluble vitamins), fat as a highly purified safflower oil or soybean oil, electrolytes, trace minerals and water. As it contains fat, that component is not in solution and therefore should have no direct effect on osmoticity. However, the potential for interactions can cause some significant changes in total particle concentration and indirectly affect the osmoticity.[8]

Although easily digested, dextrose contributes more particles than most other carbohydrate sources, such as starch, and is more likely to cause osmotic diarrhea, especially with bolus feeding. Osmoticity is improved (decreased) in the above formula by replacing dextrose with dextrose oligosaccharides (carbohydrates that yield on hydrolysis 2 to 10 monosaccharides). Flavoring also increases the osmoticity of a product, different flavors causing varying increases.

Commercial diets of this type are packaged as fluids or as powders for reconstitution. Reconstitution is usually with water. The labels of some preparations state the osmolality or osmolarity of the fluid obtained at standard dilution. However, the labels of many products do not state either their osmolality or osmolarity (or their osmoticity in any way). Often, when the term osmolarity *is* used, one cannot discern whether this is simply incorrect terminology, or whether the osmolarity actually has been calculated from the osmolality. With concentrated infant formulas or tube feedings, the osmolarity may be only 80% of the osmolality. The osmoticity (osmolality, etc) of infant formulas, tube feedings and total parenteral nutrition solutions are not described adequately either in textbooks or in the literature.

There are other areas of concern. A wide variation in osmolality was found when powdered samples from different containers were reconstituted in the same manner. This difference was found both within and among different lots of the same product. In addition, reconstitution of some powdered enteral formulas using the scoops supplied by the manufacturer gave formulas that had almost twice the osmolality of the same product when reconstituted accurately by weight.

This form of nutrition has been called, somewhat inaccurately, "Enteral Hyperalimentation."[1] It should be distinguished from (a) "Central Parenteral Nutrition" (which also has been called "Hyperalimentation," "Total Parenteral Nutrition" (TPN) and "Parenteral Hyperalimentation"); and from (b) the more recently reported "Peripheral Hyperalimentation." The terminology is in a state of flux due to the recent rapid progress in the forms of metabolic support.

The enteric route for hyperalimentation frequently is overlooked in many diseases or posttrauma states, if the patient is not readily responsive to traditional oral feedings. Poor appetite, chronic nausea, general apathy and a degree of somnolence or sedation are common concomitants of serious disease. This frequently prevents adequate oral alimentation and results in progressive energy and nutrient deficits. Often, supplementary feedings of a highly nutritious formula are taken poorly or refused entirely. However, the digestive and absorptive capabilities of the gastrointestinal tract are frequently intact and, when challenged with appropriate nutrient fluids, can be used effectively. By using an intact GI tract for proper alimentation, the major problems of sepsis and metabolic derangement which relate to intravenous hyperalimentation largely are obviated, and adequate nutritional support is simplified greatly. Because of this increased safety and ease of administration, the enteric route for hyperalimentation should be used whenever possible.[9]

When ingested in large amounts or concentrated fluids, the osmotic characteristics of certain foods can cause an upset in the normal water balance within the body. For a given weight of solute the osmolality of the solution is inversely proportional to the size of the particles. Nutritional components can be listed in an approximate order of decreasing osmotic effect per gram, as[10]

1. Electrolytes such as sodium chloride
2. Relatively small organic molecules such as dextrose (glucose) and amino acids
3. Dextrose oligosaccharides
4. Starches
5. Proteins
6. Fats (as fats are not water-soluble they have no osmotic effect)

Thus, in foods, high proportions of electrolytes, amino acids and simple sugars have the greatest effect on osmolality, and as a result, on tolerance. The approximate osmolality of a few common foods and beverages is

	mOsmol/kg
Whole milk	295
Tomato juice	595
Orange juice	935
Ice cream	1150

When nutrition of high osmoticity is ingested, large amounts of water will transfer to the stomach and intestines from the fluid surrounding those organs in an attempt to lower the osmoticity. The higher the osmoticity, the larger the amount of water required; a large amount of water in the GI tract can cause distention, cramps, nausea, vomiting, hypermotility and shock. The food may move through the tract too rapidly for the water to be reabsorbed, and result in diarrhea; severe diarrhea can cause dehydration. The hyperosmotic enteral effects have been observed by the administration of undiluted hypertonic oral medication;[11] Table 1 from this work lists average osmolality values of some commercially available drug solutions and suspensions. Thus, there is some analogy to the effect of hyperosmotic intravenous infusions.

Hyperosmotic feedings may result in mucosal damage in the GI tract. Rats given hyperosmotic feeding showed transient decrease in disaccharidase activity, and an increase in alkaline phosphatase activity. They also showed morphologic alterations in the microvilli of the small intestines. After a period of severe gastroenteritis, the bowel may be unusually susceptible to highly osmotic formulas, and their use may increase the frequency of diarrhea. Infant formulas that are hyperosmotic may affect preterm infants adversely during the early neonatal period, and they may produce or predispose neonates to necrotizing enterocolitis when delivered to the jejunum through a nasogastric tube.

The body attempts to keep the osmoticity of the contents of the stomach and intestines at approximately the same level as that of the fluid surrounding them. As a fluid of lower osmoticity requires the transfer of less water to dilute it, it should be tolerated better than one of higher osmoticity. As to tolerance, there is a great variation from one individual to another in sensitivity to the osmoticity of foods. The majority of patients receiving nutritional formulas, either orally or by tube, are able to tolerate feedings with a wide range of osmoticities if administered slowly and if adequate additional fluids are given. However, certain patients are more likely to develop symptoms of intolerance when receiving fluids of high osmoticity. These include debilitated patients, patients with GI disorders, pre- and postoperative patients, gastrostomy- and jejunostomy-fed patients and patients whose GI tracts have not been challenged for an extended period of time. Thus, osmoticity should always be considered in the selection of the formula for each individual patient. With all products, additional fluid intake may be indicated for individuals with certain clinical conditions. Frequent feedings of small volume or a continual instillation (pumped) may be of benefit initially in establishing tolerance to a formula. For other than isoosmotic formulas, feedings of reduced concentration (osmolality less than 400 mOsmol/kg) also may be helpful initially if tolerance problems arise in sensitive individuals. Concentration and size of feeding then can be increased gradually to normal as tolerance is established.

A common disturbance of intake encountered in elderly individuals relates to excess solid intake rather than to reduced water intake. For example, an elderly victim of a

cerebral vascular accident who is being fed by nasogastric tube may be given a formula whose solute load requires a greatly increased water intake. Thus, tube feeding containing 120 g of protein and 10 g of salt will result in the excretion of more than 1000 mOsmol of solute. This requires the obligatory excretion of a volume of urine between 1200 and 1500 mL when the kidneys are capable of normal concentration ability. As elderly individuals often have significant impairment in renal function, water loss as urine may exceed 2000 to 2500 mL per day. Such an individual would require 3 to 4 L of water per day simply to meet the increased demand created by this high solute intake. Failure of the physician to provide such a patient with the increased water intake needed will result in a progressive water deficit which rapidly may become critical. The importance of knowing the complete composition of the tube feeding formulas used for incapacitated patients cannot be overemphasized.

Osmolality Determination

The need for experimental determination of osmolality has been established. In regard to this there are four properties of solutions that depend only on the number of "particles" in the solution. They are osmotic pressure elevation, boiling point elevation, vapor pressure depression and freezing point depression. These are called colligative properties and if one of them is known, the others can be calculated from its value. Osmotic pressure elevation is the most difficult to measure satisfactorily. The boiling-point elevation may be determined but the values are rather sensitive to changes in barometric pressure. Also, for an aqueous solution the molal boiling-point elevation is considerably less than the freezing-point depression. Thus, it is less accurate than the freezing-point method. Determinations of vapor-pressure lowering have been considered to be impractical because of the elaborate apparatus required. However Zenk and Huxtable used a vapor pressure osmometer and state that it has much to recommend it for most of the systems under consideration here.[3] A vapor-pressure osmometer with a precision of <2 mOsmol/kg is reported by Dickerson, *et al.*[11] The method usually used is that of freezing-point depression, which can be determined quite readily with a fair degree of accuracy (see *Freezing-Point Depression*, Chapter 16). It should be noted that the data in Appendix A can be converted readily to vapor pressure lowering if desired.

Semiautomatic, high sensitivity osmometers which measure freezing point depression provide digital readouts or computer printouts of the results expressed in milliosmol units.

The results of investigations by Lund *et al*[12] indicate that the freezing point of normal, healthy human blood is −0.52° and not −0.56°, as previously assumed (see *Reliability of Data*, page 1489). Inasmuch as water is the medium in which the various constituents of blood are either suspended or dissolved in this method, it is assumed that *any aqueous solution* freezing at −0.52° is *isotonic with blood*. Now it is rare that a simple aqueous solution of the therapeutic agent to be injected parenterally has a freezing point of −0.52°, and to obtain this freezing point it is necessary either to add some other therapeutically inactive solute if the solution is hypotonic (freezing point above −0.52°) or to dilute the solution if it is hypertonic (freezing point below −0.52°). The usual practice is to add either sodium chloride or dextrose to adjust hypotonic parenteral solutions to isotonicity. Certain solutes, including ammonium chloride, boric acid, urea, glycerin and propylene glycol, cause hemolysis even when they are present in a concentration that is isoosmotic and such solutions obviously are not isotonic. See Appendix A.

In a similar manner solutions intended for ophthalmic use may be adjusted to have a freezing point identical to that of lacrimal fluid, namely, −0.52° (see *Reliability of Data*, page 1489). Ophthalmic solutions with higher freezing points usually are made isotonic by the addition of boric acid or sodium chloride.

In laboratories where the necessary equipment is available, the method usually followed for adjusting hypotonic solutions is to determine the freezing-point depression produced by the ingredients of a given prescription or formula, and then to add a quantity of a suitable inert solute calculated to lower the freezing point to −0.52°, whether the solution is for parenteral injection or ophthalmic application. A final determination of the freezing-point depression may be made to verify the accuracy of the calculation. If the solution is hypertonic, it must be diluted if an isotonic solution is to be prepared, but it must be remembered that some solutions cannot be diluted without impairing their therapeutic activity. For example, solutions to be used for treating varicose veins require a high concentration of the active ingredient (solute) to make the solution effective. Dilution to isotonic concentration is not indicated in such cases.

Freezing-Point Calculations

As explained in the preceding section, freezing-point data often may be employed in solving problems of isotonicity adjustment. Obviously, the utility of such data is limited to those solutions where the solute does not penetrate the membrane of the tissue, eg, red blood cells, with which it is in contact. In such cases, Appendix A, giving the freezing-point depression of solutions of different concentrations of various substances, provides information essential for solving the problem.

For most substances listed in the table the concentration of an isotonic solution, ie, one that has a freezing point of −0.52°, is given. If this is not listed in the table, it may be determined with sufficient accuracy by simple proportion using, as the basis for calculation, that figure which most nearly produces an isotonic solution. Actually the depression of the freezing point of a solution of an electrolyte is not absolutely proportional to the concentration but varies according to dilution; for example, a solution containing 1 g of procaine hydrochloride in 100 mL has a freezing-point depression of 0.12°, whereas a solution containing 3 g of the same salt in 100 mL has a freezing-point depression of 0.33°, *not* 0.36° (3 × 0.12°). Since the adjustment to isotonicity need not be absolutely exact, approximations may be made. When it is recalled that for many years an 0.85% solution of sodium chloride, rather than the presently employed 0.90% concentration, was accepted widely and proved to be eminently satisfactory as the isotonic equivalent of blood serum, it is apparent that minor deviations are not of great concern. Also, formerly a 1.4% solution of sodium chloride was considered to be isotonic with lacrimal fluid and found to be relatively tolerable when applied to the eye. Nevertheless, adjustments to isotonicity should be as exact as practicable.

As a specific illustration of the manner in which the data in the table may be used, suppose it is required to calculate the quantity of sodium chloride needed to make 100 mL of a 1% solution of calcium disodium edetate isoosmotic with blood serum. Reference to the table indicates that the 1% solution provides for 0.12° of the necessary 0.52° of freezing-point depression required of an isoosmotic solution, thus leaving 0.40° to be supplied by the sodium chloride. Again, referring to the table, 0.52° is found to be the freezing-point depression of a 0.9% solution of sodium chloride and by simple proportion it is calculated that a 0.69% solution will have a freezing-point depression of 0.40°. Assuming addi-

tivity of the freezing-point depressions, a solution of 0.69 g of sodium chloride and 1 g of calcium disodium edetate in sufficient water to make 100 mL will be isoosmotic with blood serum.

Likewise, to render a 1% solution of boric acid isotonic with lacrimal fluid by the addition of sodium chloride, one would proceed with the calculation as follows

Freezing-point depression of lacrimal fluid	0.52°
Freezing-point depression of 1% boric acid solution	0.29°
Freezing-point depression to be supplied by sodium chloride	0.23°
Freezing-point depression of a 0.9% solution of sodium chloride	0.52°

Therefore,

$$0.52{:}0.9 = 0.23{:}x$$
$$0.52x = 0.207$$

x = 0.4% sodium chloride to be incorporated with 1% boric acid to produce a solution which will be isotonic with lacrimal fluid.

Similarly, should a solution contain more than one ingredient, the sum of the respective freezing points of each ingredient would be determined and the difference between this sum and the required freezing point would represent the freezing point to be supplied by the added substance.

The preceding calculation can be expressed in the form of an equation, as follows

$$x = \frac{(0.52 - a) \times c}{b}$$

where

- x = g of adjusting solute required for each 100 mL of solution.
- 0.52 = Freezing point depression of blood serum or lacrimal fluid (in degrees).
- a = Freezing point depression of given ingredients in 100 mL of solution.
- b = Freezing point depression of c g of adjusting substance per 100 mL.
- c = g of adjusting solute per 100 mL, producing a freezing point depression of b.

***L*-Values**—In dilute solutions, the expression for freezing-point depression may be written as

$$\Delta T_f = Lc$$

in which ΔT_f is the freezing-point depression in °C, L is a constant and c is the molar concentration of the drug. The term, L_{iso}, is defined as the specific value of L at a concentration of drug which is isotonic with blood or lacrimal fluid.

For a more complete discussion of the use of L values, the reader is referred to RPS-14, page 1560.

Effect of Solvents—Besides water, certain other solvents frequently are employed in nose drops, ear drops and other preparations to be used in various parts of the body. Liquids such as glycerin, propylene glycol or alcohol may compose part of the solvent. In solving isotonicity adjustment problems for such solutions it should be kept in mind that while these solvent components contribute to the freezing-point depression they may or may not have an effect on the "tone" of the tissue to which they are applied, ie, an *isoosmotic* solution may not be *isotonic*. It is apparent that, in such cases, the utility of the methods described above or, for that matter, of any other method of evaluating "tonicity" is questionable.

Reliability of Data—While the freezing point of blood formerly was assumed to be −0.56°, later investigators[12] reported that as a consequence of ice being disengaged in freezing-point determinations, as ordinarily performed, the observed freezing point of blood is too low and the correct freezing point is −0.52°. The same investigators found the freezing point of a 0.9% solution of sodium chloride to be correspondingly low; the correct freezing point in this case is also −0.52°. Presumably, all solutions commonly considered to be isotonic with blood will freeze, when a correction for disengaged ice is applied, at −0.52°. It is apparent, therefore, that there is no need to change the isotonic concentration, if the reference temperature for both blood and the solution under consideration is always the same, and provided that the *method* of determining the freezing point is the same. Also, there appears to be no objection to using freezing-point data for solutions of other than isotonic concentration, if the method of determining the freezing point is the same in all cases, since any differences obtained when another method is used (such as that of Lund *et al*[12]), probably will be proportional to concentration.

In a discussion of the significance of freezing point data it is to be noted that there are some discrepancies in the literature concerning freezing points of solutions. An *exact* determination of freezing point is actually a difficult experiment; one which calls for the control of several variables which commonly are neglected, such as the disengagement of ice. It is not possible, at this time, to select unequivocal freezing point data for most of the solutions listed in Appendix A at the end of this chapter. The comprehensive and valuable data of Lund, *et al*,[12] referred to above, actually represent, in most instances, measurements of vapor pressure which have been *calculated* to corresponding freezing point depressions. It would seem to be desirable to have confirmatory evidence based on actual measurements of freezing point, determined more accurately than generally has been the case, before revisions of existing data are made. In the case of boric acid, which enters into the composition of many collyria, there is the further variable that a sterilized solution freezes at a higher temperature than a freshly prepared, unsterilized solution of the same strength. Specifically, a freshly prepared solution containing 2.85% of boric acid was found to freeze at the same temperature (−0.82°) as a 3.1% solution which had been sterilized under pressure.

Earlier in this section it was stated that at one time lacrimal fluid was considered to have the same osmotic pressure as a 1.4% solution of sodium chloride, the freezing point of which was found to be, by the usual method of determination, −0.80°. The experiments of Krogh, *et al*[13] have indicated that lacrimal fluid has the same osmotic pressure as blood and, that instead of assuming that the freezing point of solutions isotonic with lacrimal fluid is −0.80°, it should be the same as that of blood, namely, −0.52°. Accordingly, the procedure for adjusting solutions to isotonicity with lacrimal fluid is qualitatively and quantitatively the same as the procedure for blood.

Tonicity Testing by Observing Erythrocyte Changes

Observation of the behavior of human erythrocytes when suspended in a solution is the ultimate and direct procedure for determining whether the solution is isotonic, hypotonic or hypertonic. If hemolysis or marked change in the appearance of the erythrocytes occurs, the solution is not isotonic with the cells. If the cells retain their normal characteristics, the solution is isotonic.

Hemolysis may occur when the osmotic pressure of the fluid in the erythrocytes is greater than that of the solution in which the cells are suspended, but the specific chemical reactivity of the solute in the solution often is far more important in producing hemolysis than is the osmotic effect. There is no certain evidence that any single mechanism of action causes hemolysis. The process appears to involve such factors as pH, lipid solubility, molecular and ionic sizes of solute particles and possibly inhibition of cholinesterase

in cell membranes and denaturing action on plasma membrane protein.

Some investigators test the tonicity of injectable solutions by observing variations of red-blood-cell volume produced by these solutions. This method appears to be more sensitive to small differences in tonicity than those based on observation of a hemolytic effect. Much useful information concerning the effect of various solutes on erythrocytes has been obtained by this procedure and a summary of many of these data is given in RPS-14, page 1562.

Other Methods of Adjusting Tonicity

Several methods for adjusting tonicity, other than those already described, are used.

Sodium Chloride Equivalent Methods—A sodium chloride equivalent is defined as the weight of sodium chloride which will produce the same osmotic effect as 1 g of the drug prepared as an isotonic solution. Appendix A lists the sodium chloride equivalents for many drugs. Some of the equivalents vary with the concentration of the drug (in certain cases because of changes of interionic attraction at different concentrations) but, in every case, the equivalent is for 1 g of drug. As an example of the use of these data, if the sodium chloride equivalent of boric acid is 0.5 at 1% concentration, this is interpreted to mean that 1 g of boric acid in solution will produce the same freezing-point depression as 0.5 g of sodium chloride, or that a 1% boric acid solution is equivalent in its colligative properties to a 0.5% solution of sodium chloride. From Appendix A it is found that for a 1.9% boric acid solution (ie, at isotonicity) the sodium chloride equivalent is 0.47, corresponding to a 0.9% sodium chloride solution (1.9×0.47).

Examples illustrating use of the sodium chloride equivalent method to adjust collyria to isotonicity follow. The same type of calculation may be used for other solutions that are to be made isotonic.

Example 1

Homatropine Hydrobromide 1%
to make collyr isotonic ..... 60 mL

0.6 g of homatropine hydrobromide is required. 1 g or 1% of the drug is equivalent in osmotic effect to 0.17 g or 0.17% of sodium chloride.

$$0.17 \times 0.6 = 0.102 \text{ g (sodium chloride)}$$

60 mL of an isotonic sodium chloride solution contains 0.54 g sodium chloride
0.6 g homatropine hydrobromide is equivalent to 0.102 g sodium chloride
0.438 g sodium chloride

Therefore, 0.438 g of sodium chloride must be added to make 60 mL of a 1% homatropine hydrobromide solution isotonic with tear fluid. The same calculations may be made using percentage calculations. 1% of homatropine hydrobromide corresponds to 0.17% sodium chloride in colligative properties.

Thus, 0.9% − 0.17% = 0.73% must be added, 0.73% of 60 mL = 0.438 g of sodium chloride to be added.

If boric acid is to be used as the adjusting substance the calculations have to be carried one step further. There is no "boric acid equivalent," but the sodium chloride equivalent of boric acid at 1% concentration is 0.5, meaning that 1 g of boric acid (or 1%) corresponds in colligative properties to 0.5 g sodium chloride (or 0.5%). Using the result obtained above, which was 0.438 g of sodium chloride to be added, it now follows that the sodium chloride equivalent of boric acid must be divided into the amount of sodium chloride or expressed as an equation:

$$1 \text{ g boric acid} : 0.5 \text{ g sodium chloride} = x \text{ g} : 0.438 \text{ g}$$
$$x = 0.876 \text{ g boric acid to be added}$$

For a prescription containing more than one active drug, the calculations for sodium chloride are carried out separately, the obtained quantities are added, and then the total is deducted from the 0.9% amount.

Example 2

Epinephrine Hydrochloride 0.5%
Zinc Sulfate 0.3%
Sterile Water qs, to make ..... 30 mL
M Ft Collyr isotonic SA

Sodium chloride equivalent of epinephrine HCl is 0.29
Sodium chloride equivalent of zinc sulfate is 0.15

150 mg epinephrine hydrochloride	~43.5 mg sodium chloride
90 mg zinc sulfate	~13.5 mg sodium chloride
Total ingredients are equivalent to	~57 mg sodium chloride

0.9% of 30 mL 270 mg sodium chloride
57 mg
213 mg

213 mg of sodium chloride must be added to make this solution isotonic with tear fluid. Since boric acid is the adjusting substance of choice for the solution 426 mg should be used (0.5 divided into 213 mg).

Isotonic Solution V-Values—These are the volumes of sterile water to be added to a specified weight of drug (often 0.3 g but sometimes 1 g) to prepare an isotonic solution. Appendix B gives such values for some commonly used drugs. The reason for providing data for 0.3 g drug is for the convenience of preparing 30 mL (1 fl oz) of solution, as is prescribed often. If values for 100 mL of final solution are desired, the data in Appendix B should be multiplied by 100/30. The basic principle underlying the use of these values is to prepare an isotonic solution of the prescribed drug in sterile water and then dilute this solution to the required final volume with a suitable isotonic vehicle. For example, if 0.3 g of a drug is specified to be used (as in preparing 30 mL of 1% solution of the drug), it is first dissolved in the volume of sterile water stated in Appendix B and then diluted to 30 mL with a suitable isotonic vehicle. Isotonic solution values can be used, of course, for calculating tonicity-adjusting data for concentrations of drugs other than 1% and for volumes other than 30 mL. How this is done is illustrated in the following examples.

Example 1

A prescription calls for

Atropine sulfate ..... 0.3 g
Sterile water qs ..... 60 mL

M Ft Collyr isotonic and buffered SA
Sig: For Office Use.

This order is for a 0.5% solution of atropine sulfate. According to Appendix B, 0.3 g of atropine sulfate dissolved in 4.3 mL of sterile water will produce a 1% isotonic solution when diluted to 30 mL with an isotonic vehicle. For 30 mL of 0.5% solution, half the quantities of atropine sulfate and sterile water would be used, but for 60 mL of 0.5% solution the same quantities as for 30 mL of 1% solution are required.

Therefore, to fill this prescription order, 0.3 g of atropine sulfate should be dissolved in 4.3 mL of sterile preserved water and diluted with isotonic preserved Sørensen's pH 6.8 phosphate buffer to 60 mL.

* * * *

For more than one active ingredient in solution the quantity of water to be used is calculated separately for each ingredient. The values thus obtained are added, the total amount of sterile water then is used to dissolve the active ingredients and finally sufficient isotonic, buffered preserved solution (diluting solution) is used to make the required volume.

Example 2

A prescription calls for

Epinephrine hydrochloride 0.5%
Zinc sulfate 0.3%
Sterile water qs to make ..... 30 mL

M Ft Collyr isotonic

In this example the active ingredients are given in percentage. The ideal vehicle is 1.9% boric acid solution. Reference to the table for isotonic solution values shows the following.

Epinephrine hydrochloride 0.3 g (1%) will make 9.7 mL of an isotonic solution when dissolved in sterile preserved water. Zinc sulfate 0.3 g will make 5 mL of an isotonic solution with sterile water.

Therefore, the quantities called for in this prescription will make 4.85 mL and 1.5 mL of isotonic solutions, respectively. Dissolve the salts in sufficient sterile preserved water to make 6.35 mL and add sufficient 1.9% preserved boric acid solution to make 30 mL. The resulting solution is isotonic.

Since it is practically impossible to measure the required volumes accurately, it is feasible, in this instance, to use 6.35 mL of sterile water as the total solvent for these two drugs. Graduated pipets, previously sterilized, are necessary for this work.

References

1. Kaminski MV: *Surg Gyn Obst 143:* 12, 1976.
2. Theeuwes F: *J Pharm Sci 64:* 1987, 1975.
3. Zenk K, Huxtable RF: *Hosp Pharm 13:* 577, 1978.
4. McDuffee L: *IL Council Hosp Pharm Drug Inf Newsl 8:* Oct.–Nov. 1978.
5. Streng WH, *et al.: J Pharm Sci 67:* 394, 1978.
6. Murty BSR, *et al: Am J Hosp Pharm 33:* 546, 1976.
7. Bray AJ: Personal Communication, Mead Johnson Nutritional Division, Evansville IN, 1978.
8. Andrassy RJ, *et al: Surg 82:* 205, 1977.
9. Dobbie RP, Hoffmeister JA: *Surg Gyn Obst 143:* 273, 1976.
10. Anon. Osmolality: Doyle Pharmaceutical Co, Minneapolis, 1978.
11. Dickerson RN, Mehnik G, *Am J Hosp Pharm 45:* 832, 1988. Holtz L, Milton J, Sturek JK: *J Parent Ent Nutr 11:* 183, 1987.
12. Lund CG, *et al: The Preparation of Solutions Iso-osmotic with Blood, Tears, and Tissue.* Danish Pharmacopoeial Commission, Einar Munksgaard, Copenhagen, 1947.
13. Krogh A, *et al: Acta Physiol Scand 10:* 88, 1945.
14. Hammarlund ER, *et al: J Pharm Sci 54:* 160, 1965.
15. Hammarlund ER, Pedersen-Bjergaard K: *J APhA Sci Ed 47:* 107, 1958.
16. Hammarlund ER, Pedersen-Bjergaard K: *J Pharm Sci 50:* 24, 1961.
17. Hammarlund ER, Van Pevenage GL: *J Pharm Sci 55:* 1448, 1966.
18. Sapp C, *et al: J Pharm Sci 64:* 1884, 1975.
19. *British Pharmaceutical Codex,* Pharmaceutical Press, London, 1973.
20. Fassett WE, *et al: J Pharm Sci 58:* 1540, 1969.
21. Kogan DG, Kinsey VE: *Arch Ophthalmol 27:* 696, 1942.

Bibliography

Alberty RA, Daniels F: *Physical Chemistry,* 7th ed, Wiley, New York, 1988.

Anthony CP, Kolthoff NJ: *Textbook of Anatomy and Physiology,* Mosby, St Louis, 1971.

Cowan G, Scheetz W, eds: *Intravenous Hyperalimentation,* Lea & Febiger, Philadelphia, 1972.

Garb S: *Laboratory Tests,* 6th ed, Springer, New York, 1976.

Hall WE: *Am J Pharm Educ 34:* 204, 1970.

Harvey AM, Johns RJ, Owens AH, Ross RS: *The Principles and Practice of Medicine,* 18th ed, Appleton-Century-Crofts, New York, 1972.

Martin AN, Swarbrick J, Cammarata A: *Physical Pharmacy,* 3rd ed, Lea & Febiger, Philadelphia, 1983.

Plumer AL: *Principles and Practice of Intravenous Therapy,* 4th ed, Little, Brown, Boston, 1987.

Ravel R: *Clinical Laboratory Medicine,* Year Book, Chicago, 1988.

Tilkian SM, Conover MH: *Clinical Implications of Laboratory Tests,* 4th ed, Mosby, St Louis, 1987.

Turco S, King RE: *Sterile Dosage Forms,* 3rd ed, Lea & Febiger, Philadelphia, 1987.

Wallach J: *Interpretation of Diagnostic Tests,* 4th ed, Little, Brown, Boston, 1986.

Appendix A—Sodium Chloride Equivalents, Freezing-Point Depressions and Hemolytic Effects of Certain Medicinals in Aqueous Solution

	0.5%		1%		2%		3%		5%		Isoosmotic concentration[a]				
	E	D	E	D	E	D	E	D	E	D	%	E	D	H	pH
Acetrizoate methylglucamine	0.09		0.08		0.08		0.08		0.08		12.12	0.07		0	7.1
Acetrizoate sodium	0.10	0.027	0.10	0.055	0.10	0.109	0.10	0.163	0.10	0.273	9.64	0.09	0.52	0	6.9[j]
Acetylcysteine	0.20	0.056	0.20	0.113	0.20	0.227	0.20	0.341			4.58	0.20	0.52	100[k]	2.0
Adrenaline HCl											4.24			68	4.5
Alphaprodine HCl	0.19	0.053	0.19	0.105	0.18	0.212	0.18	0.315			4.98	0.18	0.52	100	5.3
Alum (potassium)			0.18				0.15		0.15		6.36	0.14		24[a]	3.4
Amantadine HCl	0.31	0.090	0.31	0.180	0.31	0.354					2.95	0.31	0.52	91	5.7
Aminoacetic acid	0.42	0.119	0.41	0.235	0.41	0.470					2.20	0.41	0.52	0[a]	6.2
Aminohippuric acid	0.13	0.035	0.13	0.075											
Aminophylline				0.098[c]											
Ammonium carbonate	0.70	0.202	0.70	0.405							1.29	0.70	0.52	97	7.7
Ammonium chloride			1.12								0.8	1.12	0.52	93	5.0
Ammonium lactate	0.33	0.093	0.33	0.185	0.33	0.370					2.76	0.33	0.52	98	5.9
Ammonium nitrate	0.69	0.200	0.69	0.400							1.30	0.69	0.52	91	5.3
Ammonium phosphate, dibasic	0.58	0.165	0.55	0.315							1.76	0.51	0.52	0	7.9
Ammonium sulfate	0.55	0.158	0.55	0.315							1.68	0.54	0.52	0	5.3
Amobarbital sodium			0.25	0.143[c]			0.25				3.6	0.25	0.52	0	9.3
d-Amphetamine HCl											2.64			98	5.7
Amphetamine phosphate			0.34	0.20			0.27	0.47			3.47	0.26	0.52	0	4.5
Amphetamine sulfate			0.22	0.129[c]			0.21	0.36			4.23	0.21	0.52	0	5.9
Amprotropine phosphate											5.90			0	4.2
Amylcaine HCl			0.22				0.19				4.98	0.18		100	5.6
Aniferidine HCl	0.19	0.052	0.19	0.104	0.19	0.212	0.18	0.316	0.18	0.509	5.13	0.18	0.52	12	2.6
Antazoline phosphate											6.05			90	4.0
Antimony potassium tartrate			0.18				0.13		0.10						
Antipyrine			0.17	0.10			0.14	0.24	0.14	0.40	6.81	0.13	0.52	100	6.1
Apomorphine HCl			0.14	0.080[c]											
Arginine glutamate	0.17	0.048	0.17	0.097	0.17	0.195	0.17	0.292	0.17	0.487	5.37	0.17	0.52	0	6.9
Ascorbic acid				0.105[c]							5.05		0.52[b]	100[a]	2.2
Atropine methylbromide			0.14				0.13		0.13		7.03	0.13			
Atropine methylnitrate											6.52			0	5.2
Atropine sulfate			0.13	0.075			0.11	0.19	0.11	0.32	8.85	0.10	0.52	0	5.0
Bacitracin			0.05	0.03			0.04	0.07	0.04	0.12					
Barbital sodium			0.30	0.171[c]			0.29	0.50			3.12	0.29	0.52	0	9.6
Benzalkonium chloride			0.16				0.14		0.13						
Benztropine mesylate	0.26	0.073	0.21	0.115	0.15	0.170	0.12	0.203	0.09	0.242					
Benzyl alcohol			0.17	0.09[c]			0.15								

Appendix A—Continued

	0.5%		1%		2%		3%		5%		Isoosmotic concentration[a]				
	E	D	E	D	E	D	E	D	E	D	%	E	D	H	pH
Bethanechol chloride	0.50	0.140	0.39	0.225	0.32	0.368	0.30	0.512			3.05	0.30		0	6.0
Bismuth potassium tartrate			0.09				0.06		0.05						
Bismuth sodium tartrate			0.13				0.12		0.11		8.91	0.10		0	6.1
Boric acid			0.50	0.288[c]							1.9	0.47	0.52	100	4.6
Brompheniramine maleate	0.10	0.026	0.09	0.050	0.08	0.084									
Bupivacaine HCl	0.17	0.048	0.17	0.096	0.17	0.193	0.17	0.290	0.17	0.484	5.38	0.17	0.52	83	6.8
Butabarbital sodium	0.27	0.078	0.27	0.155	0.27	0.313	0.27	0.470			3.33	0.27	0.52	0	6.8
Butacaine sulfate			0.20	0.12			0.13	0.23	0.10	0.29					
Caffeine and sodium benzoate			0.26	0.15			0.23	0.40			3.92	0.23	0.52	0	7.0
Caffeine and sodium salicylate			0.12	0.12			0.17	0.295	0.16	0.46	5.77	0.16	0.52	0	6.8
Calcium aminosalicylate											4.80			0	6.0
Calcium chloride			0.51	0.298[c]							1.70	0.53	0.52	0	5.6
Calcium chloride ($6\ H_2O$)			0.35	0.20							2.5	0.30	0.52	0	5.7
Calcium chloride, anhydrous			0.68	0.39							1.3	0.69	0.52	0	5.6
Calcium disodium edetate	0.21	0.061	0.21	0.120	0.21	0.240	0.20	0.357			1.50	0.20	0.52	0	6.1
Calcium gluconate			0.16	0.091[c]			0.14	0.24							
Calcium lactate			0.23	0.13			0.12	0.36			4.5	0.20	0.52	0	6.7
Calcium lactobionate	0.08	0.022	0.08	0.043	0.08	0.085	0.07	0.126	0.07	0.197					
Calcium levulinate			0.27	0.16			0.25	0.43			3.58			0	7.2
Calcium pantothenate											5.50			0	7.4
Camphor				0.12[d]											
Capreomycin sulfate	0.04	0.011	0.04	0.020	0.04	0.042	0.04	0.063	0.04	0.106					
Carbachol				0.205[c]							2.82			0	5.9
Carbenicillin sodium	0.20	0.059	0.20	0.118	0.20	0.236	0.20	0.356			4.40	0.20	0.52	0	6.6
Carboxymethylcellulose sodium	0.03	0.007	0.03	0.017											
Cephaloridine	0.09	0.023	0.07	0.041	0.06	0.074	0.06	0.106	0.05	0.145					
Chloramine-T											4.10			100*	9.1
Chloramphenicol				0.06[d]											
Chloramphenicol sodium succinate	0.14	0.038	0.14	0.078	0.14	0.154	0.13	0.230	0.13	0.382	6.83	0.13	0.52	partial	6.1
Chlordiazepoxide HCl	0.24	0.068	0.22	0.125	0.19	0.220	0.18	0.316	0.17	0.487	5.50	0.16	0.52	66	2.7
Chlorobutanol (hydrated)			0.24	0.14											
Chloroprocaine HCl	0.20	0.054	0.20	0.108	0.18	0.210									
Chloroquine phosphate	0.14	0.039	0.14	0.082	0.14	0.162	0.14	0.242	0.13	0.379	7.15	0.13	0.52	0	4.3
Chloroquine sulfate	0.10	0.028	0.09	0.050	0.08	0.090	0.07	0.127	0.07	0.195					
Chlorpheniramine maleate	0.17	0.048	0.15	0.085	0.14	0.166	0.13	0.220	0.09	0.265					
Chlortetracycline HCl	0.10	0.030	0.10	0.061	0.10	0.121									
Chlortetracycline sulfate			0.13	0.08			0.10	0.17							
Citric acid			0.18	0.10			0.17	0.295	0.16	0.46	5.52	0.16	0.52	100*	1.8
Clindamycin phosphate	0.08	0.022	0.08	0.046	0.08	0.095	0.08	0.144	0.08	0.242	10.73	0.08	0.52	58*	6.8
Cocaine HCl			0.16	0.090[c]			0.15	0.26	0.14	0.40	6.33	0.14	0.52	47	4.4
Codeine phosphate			0.14	0.080[c]			0.13	0.23	0.13	0.38	7.29	0.12	0.52	0	4.4
Colistimethate sodium	0.15	0.045	0.15	0.085	0.15	0.170	0.15	0.253	0.14	0.411	6.73	0.13	0.52	0	7.6
Cupric sulfate			0.18	0.100[c]			0.15		0.14		6.85	0.13		trace*	3.9
Cyclizine HCl	0.20	0.060													
Cyclophosphamide	0.10	0.031	0.10	0.061	0.10	0.125									
Cytarabine	0.11	0.034	0.11	0.066	0.11	0.134	0.11	0.198	0.11	0.317	8.92	0.10	0.52	0	8.0
Deferoxamine mesylate	0.09	0.023	0.09	0.047	0.09	0.093	0.09	0.142	0.09	0.241					
Demecarium bromide	0.14	0.038	0.12	0.069	0.10	0.108	0.08	0.139	0.07	0.192					
Dexamethasone sodium phosphate	0.18	0.050	0.17	0.095	0.16	0.180	0.15	0.260	0.14	0.410	6.75	0.13	0.52	0	8.9
Dextroamphetamine HCl	0.34	0.097	0.34	0.196	0.34	0.392					2.64	0.34	0.52		
Dextroamphetamine phosphate			0.25	0.14			0.25	0.44			3.62	0.25	0.52	0	4.7
Dextroamphetamine sulfate	0.24	0.069	0.23	0.134	0.22	0.259	0.22	0.380			4.16	0.22	0.52	0	5.9
Dextrose			0.16	0.091[c]			0.16	0.28	0.16	0.46	5.51	0.16	0.52	0	5.9
Dextrose (anhydrous)			0.18	0.101[c]			0.18	0.31			5.05	0.18	0.52	0	6.0
Diatrizoate sodium	0.10	0.025	0.09	0.049	0.09	0.098	0.09	0.149	0.09	0.248	10.55	0.09	0.52	0	7.9
Dibucaine HCl				0.074[c]											
Dicloxacillin sodium ($1\ H_2O$)	0.10	0.030	0.10	0.061	0.10	0.122	0.10	0.182							

Appendix A—Continued

	0.5%		1%		2%		3%		5%		Isoosmotic concentration[a]				
	E	D	E	D	E	D	E	D	E	D	%	E	D	H	pH
Diethanolamine	0.31	0.089	0.31	0.177	0.31	0.358					2.90	0.31	0.52	100	11.3
Dihydrostreptomycin sulfate			0.06	0.03			0.05	0.09	0.05	0.14	19.4	0.05	0.52	0	6.1
Dimethpyrindene maleate	0.13	0.039	0.12	0.070	0.11	0.120									
Dimethyl sulfoxide	0.42	0.122	0.42	0.245	0.42	0.480					2.16	0.42	0.52	100	7.6
Diperodon HCl	0.15	0.045	0.14	0.079	0.13	0.141									
Diphenhydramine HCl				0.161[c]							5.70			88*	5.5
Diphenidol HCl	0.16	0.045	0.16	0.09	0.16	0.180									
Doxapram HCl	0.12	0.036	0.12	0.070	0.12	0.140	0.12	0.210							
Doxycycline hyclate	0.12	0.036	0.12	0.072	0.12	0.134	0.11	0.186	0.09	0.264					
Dyphylline	0.10	0.025	0.10	0.052	0.09	0.104	0.09	0.155	0.08	0.245					
Echothiophate iodide	0.16	0.045	0.16	0.090	0.16	0.170									
Edetate disodium	0.24	0.070	0.23	0.132	0.22	0.248	0.21	0.360			4.44	0.20	0.52	0	4.7
Edetate trisodium monohydrate	0.29	0.079	0.29	0.158	0.28	0.316	0.27	0.472			3.31	0.27	0.52	0	8.0
Emetine HCl				0.058[c]				0.17		0.29					
Ephedrine HCl			0.30	0.165[c]			0.28				3.2	0.28		96	5.9
Ephedrine sulfate			0.23	0.13			0.20	0.36			4.54	0.20	0.52	0	5.7
Epinephrine bitartrate			0.18	0.104			0.16	0.28	0.16	0.462	5.7	0.16	0.52	100*	3.4
Epinephrine hydrochloride			0.29	0.16[b]			0.26				3.47	0.26			
Ergonovine maleate				0.089[c]											
Erythromycin lactobionate	0.08	0.020	0.07	0.040	0.07	0.078	0.07	0.115	0.06	0.187					
Ethyl alcohol											1.39			100	6.0
Ethylenediamine				0.253[c]							2.08			100*	11.4
Ethylmorphine HCl			0.16	0.088[c]			0.15	0.26	0.15	0.43	6.18	0.15	0.52	38	4.7
Eucatropine HCl				0.11[d]											
Ferric ammonium citrate (green)											6.83			0	5.2
Floxuridine	0.14	0.040	0.13	0.076	0.13	0.147	0.12	0.213	0.12	0.335	8.47	0.12	0.52	3*	4.5
Fluorescein sodium			0.31	0.181[c]			0.27	0.47			3.34	0.27	0.52	0	8.7
Fluphenazine 2-HCl	0.14	0.041	0.14	0.082	0.12	0.145	0.09	0.155							
d-Fructose											5.05			0*	5.9
Furtrethonium iodide	0.24	0.070	0.24	0.133	0.22	0.250	0.21	0.360			4.44	0.20	0.52	0	5.4
Galactose											4.92			0	5.9
Gentamicin sulfate	0.05	0.015	0.05	0.030	0.05	0.060	0.05	0.093	0.05	0.153					
D-Glucuronic acid											5.02			48*	1.6
Glycerin				0.203[c]							2.6			100	5.9
Glycopyrrolate	0.16	0.042	0.15	0.084	0.15	0.166	0.14	0.242	0.13	0.381	7.22	0.12	0.52	92*	4.0
Gold sodium thiomalate	0.10	0.032	0.10	0.061	0.10	0.111	0.09	0.159	0.09	0.250					
Hetacillin potassium	0.17	0.048	0.17	0.095	0.17	0.190	0.17	0.284	0.17	0.474	5.50	0.17	0.52	0	6.3
Hexafluorenium bromide	0.12	0.033	0.11	0.065											
Hexamethonium tartrate	0.16	0.045	0.16	0.089	0.16	0.181	0.16	0.271	0.16	0.456	5.68	0.16	0.52		
Hexamethylene sodium acetaminosalicylate	0.18	0.049	0.18	0.099	0.17	0.199	0.17	0.297	0.16	0.485	5.48	0.16	0.52	0*	4.0
Hexobarbital sodium				0.15[c]											
Hexylcaine HCl											4.30			100	4.8
Histamine 2HCl	0.40	0.115	0.40	0.233	0.40	0.466					2.24	0.40	0.52	79*	3.7
Histamine phosphate				0.149[c]							4.10			0	4.6
Histidine HCl											3.45			40	3.9
Holocaine HCl			0.20	0.12											
Homatropine hydrobromide			0.17	0.097[c]			0.16	0.28	0.16	0.46	5.67	0.16	0.52	92	5.0
Homatropine methylbromide			0.19	0.11			0.15	0.26	0.13	0.38					
4-Homosulfanilamide HCl											3.69			0	4.9
Hyaluronidase	0.01	0.004	0.01	0.007	0.01	0.013	0.01	0.020	0.01	0.033					
Hydromorphone HCl											6.39			64	5.6
Hydroxyamphetamine HBr				0.15[d]							3.71			92	5.0
8-Hydroxyquinoline sulfate											9.75			59*	2.5
Hydroxystilbamidine isethionate	0.20	0.060	0.16	0.090	0.12	0.137	0.10	0.170	0.07	0.216					
Hyoscyamine hydrobromide											6.53			68	5.9
Imipramine HCl	0.20	0.068	0.20	0.110	0.13	0.143									

Appendix A—Continued

	0.5%		1%		2%		3%		5%		Isosmotic concentration[a]				
	E	D	E	D	E	D	E	D	E	D	%	E	D	H	pH
Indigotindisulfonate sodium	0.30	0.085	0.30	0.172											
Intracaine HCl											4.97			85	5.0
Iodophthalein sodium				0.07[c]							9.58			100	9.4
Isometheptene mucate	0.18	0.048	0.18	0.095	0.18	0.196	0.18	0.302			4.95	0.18	0.52	0	6.2
Isoproterenol sulfate	0.14	0.039	0.14	0.078	0.14	0.156	0.14	0.234	0.14	0.389	6.65	0.14	0.52	trace	4.5
Kanamycin sulfate	0.08	0.021	0.07	0.041	0.07	0.083	0.07	0.125	0.07	0.210					
Lactic acid				0.230[c]							2.30			100*	2.1
Lactose			0.07	0.040[c]			0.08		0.09		9.75	0.09		0*	5.8
Levallorphan tartrate	0.13	0.036	0.13	0.073	0.13	0.143	0.12	0.210	0.12	0.329	9.40	0.10	0.52	59*	6.9
Levorphanol tartrate	0.12	0.033	0.12	0.067	0.12	0.136	0.12	0.203							
Lidocaine HCl				0.13[c]							4.42			85	4.3
Lincomycin HCl	0.16	0.045	0.16	0.090	0.15	0.170	0.14	0.247	0.14	0.400	6.60	0.14	0.52	0	4.5
Lobeline HCl				0.09[b]											
Lyapolate sodium	0.10	0.025	0.09	0.051	0.09	0.103	0.09	0.157	0.09	0.263	9.96	0.09	0.52	0	6.5[f]
Magnesium chloride				0.45							2.02	0.45		0	6.3
Magnesium sulfate			0.17	0.094[c]			0.15	0.26	0.15	0.43	6.3	0.14	0.52	0	6.2
Magnesium sulfate, anhydrous	0.34	0.093	0.32	0.184	0.30	0.345	0.29	0.496			3.18	0.28	0.52	0	7.0
Mannitol				0.098[c]						5.07			0*	6.2	
Maphenide HCl	0.27	0.075	0.27	0.153	0.27	0.303	0.26	0.448			3.55	0.25	0.52		
Menadiol sodium diphosphate											4.36			0	8.2
Menadione sodium bisulfite											5.07			0	5.3
Menthol				0.12[d]											
Meperidine HCl				0.125[c]							4.80			98	5.0
Mepivacaine HCl	0.21	0.060	0.21	0.116	0.20	0.230	0.20	0.342			4.60	0.20	0.52	45	4.5
Merbromin				0.08[b]											
Mercuric cyanide			0.15				0.14		0.13						
Mersalyl				0.06[b]											
Mesoridazine besylate	0.10	0.024	0.07	0.040	0.05	0.058	0.04	0.071	0.03	0.087					
Metaraminol bitartrate	0.20	0.060	0.20	0.112	0.19	0.210	0.18	0.308	0.17	0.505	5.17	0.17	0.52	59	3.8
Methacholine chloride				0.184[c]							3.21			0	4.5
Methadone HCl				0.101[c]							8.59			100*	5.0
Methamphetamine HCl				0.213[c]							2.75			97	5.9
Methdilazine HCl	0.12	0.035	0.10	0.056	0.08	0.080	0.06	0.093	0.04	0.112					
Methenamine			0.23				0.24				3.68	0.25		100	8.4
Methiodal sodium	0.24	0.068	0.24	0.136	0.24	0.274	0.24	0.410			3.81	0.24	0.52	0	5.0
Methitural sodium	0.26	0.074	0.25	0.142	0.24	0.275	0.23	0.407			3.85	0.23	0.52	78	9.8
Methocarbamol	0.10	0.030	0.10	0.060											
Methotrimeprazine HCl	0.12	0.034	0.10	0.060	0.07	0.077	0.06	0.094	0.04	0.125					
Methoxyphenamine HCl	0.26	0.075	0.26	0.150	0.26	0.300	0.26	0.450			3.47	0.26	0.52	96	5.4
p-Methylaminoethanolphenol tartrate	0.18	0.048	0.17	0.095	0.16	0.190	0.16	0.282	0.16	0.453	5.83	0.16	0.52	0	6.2
Methyldopate HCl	0.21	0.063	0.21	0.122	0.21	0.244	0.21	0.365			4.28	0.21	0.52	partial	3.0
Methylergonovine maleate	0.10	0.028	0.10	0.056											
N-Methylglucamine	0.20	0.057	0.20	0.111	0.18	0.214	0.18	0.315	0.18	0.517	5.02	0.18	0.52	4	11.3
Methylphenidate HCl	0.22	0.065	0.22	0.127	0.22	0.258	0.22	0.388			4.07	0.22	0.52	66	4.3
Methylprednisolone Na succinate	0.10	0.025	0.09	0.051	0.09	0.102	0.08	0.143	0.07	0.200					
Minocycline HCl	0.10	0.030	0.10	0.058	0.09	0.107	0.08	0.146							
Monoethanolamine	0.53	0.154	0.53	0.306							1.70	0.53	0.52	100	11.4
Morphine HCl			0.15	0.086[c]			0.14								
Morphine sulfate			0.14	0.079[c]			0.11	0.19	0.09	0.26					
Nalorphine HCl	0.24	0.070	0.21	0.121	0.18	0.210	0.17	0.288	0.15	0.434	6.36	0.14	0.52	63	4.1
Naloxone HCl	0.14	0.042	0.14	0.083	0.14	0.166	0.13	0.230	0.13	0.367	8.07	0.11	0.52	35	5.2
Naphazoline HCl			0.27	0.14[d]			0.24				3.00	0.22		100	5.3
Neoarsphenamine											2.32			17	7.8
Neomycin sulfate			0.11	0.063[c]			0.09	0.16	0.08	0.232					
Neostigmine bromide			0.22	0.127[c]			0.19				4.98			0	4.6
Neostigmine methylsulfate			0.20	0.115[c]			0.18		0.17		6.22	0.17			
Nicotinamide			0.26	0.148[c]			0.21	0.36			4.49	0.20	0.52	100	7.0
Nicotinic acid			0.25	0.144[c]											
Nikethamide				0.100[c]							5.94			100	6.9
Novobiocin sodium	0.12	0.033	0.10	0.057	0.07	0.073									
Oleandomycin phosphate	0.08	0.017	0.08	0.038	0.08	0.084	0.08	0.129	0.08	0.255	10.82	0.08	0.52	0	5.0
Orphenadrine citrate	0.13	0.037	0.13	0.074	0.13	0.144	0.12	0.204	0.10	0.285					
Oxophenarsine HCl											3.67			trace*	2.3

Appendix A—Continued

	0.5%		1%		2%		3%		5%		Isosmotic concentration[a]				
	E	D	E	D	E	D	E	D	E	D	%	E	D	H	pH
Oxymetazoline HCl	0.22	0.063	0.22	0.124	0.20	0.232	0.19	0.335			4.92	0.18	0.52	86	5.7
Oxyquinoline sulfate	0.24	0.068	0.23	0.113	0.16	0.182	0.14	0.236	0.11	0.315					
d-Pantothenyl alcohol	0.20	0.053	0.18	0.100	0.17	0.193	0.17	0.283	0.16	0.468	5.60	0.16	0.52	92	6.8
Papaverine HCl			0.10	0.061[c]											
Paraldehyde	0.25	0.071	0.25	0.142	0.25	0.288	0.25	0.430			3.65	0.25	0.52	97	6.3
Pargyline HCl	0.30	0.083	0.29	0.165	0.29	0.327	0.28	0.491			3.18	0.28	0.52	91	3.8
Penicillin G, potassium			0.18	0.102[c]			0.17	0.29	0.16	0.46	5.48	0.16	0.52	0	6.2
Penicillin G, procaine				0.06[d]											
Penicillin G, sodium			0.18	0.100[c]			0.16	0.28	0.16	0.46	5.90			18	5.2
Pentazocine lactate	0.15	0.042	0.15	0.085	0.15	0.169	0.15	0.253	0.15	0.420					
Pentobarbital sodium				0.145[c]							4.07			0	9.9
Pentolinium tartrate											5.95			55*	3.4
Phenacaine HCl				0.09[d]											
Pheniramine maleate				0.09[d]											
Phenobarbital sodium			0.24	0.135[c]			0.23	0.40			3.95	0.23	0.52	0	9.2
Phenol			0.35	0.20							2.8	0.32	0.52	0*	5.6
Phentolamine mesylate	0.18	0.052	0.17	0.096	0.16	0.173	0.14	0.244	0.13	0.364	8.23	0.11	0.52	83	3.5
Phenylephrine HCl			0.32	0.184[c]			0.30				3.0	0.30		0	4.5
Phenylephrine tartrate											5.90			58*	5.4
Phenylethyl alcohol	0.25	0.070	0.25	0.141	0.25	0.283									
Phenylpropanolamine HCl			0.38	0.219[c]							2.6	0.35		95	5.3
Physostigmine salicylate			0.16	0.090[c]											
Physostigmine sulfate				0.074[c]											
Pilocarpine HCl			0.24	0.138[c]			0.22	0.38			4.08	0.22	0.52	89	4.0
Pilocarpine nitrate			0.23	0.132[c]			0.20	0.35			4.84	0.20	0.52	88	3.9
Piperocaine HCl				0.12[d]							5.22			65	5.7
Polyethylene glycol 300	0.12	0.034	0.12	0.069	0.12	0.141	0.12	0.216	0.13	0.378	6.73	0.13	0.52	53	3.8
Polyethylene glycol 400	0.08	0.022	0.08	0.047	0.09	0.098	0.09	0.153	0.09	0.272	8.50	0.11	0.52	0	4.4
Polyethylene glycol 1500	0.06	0.015	0.06	0.036	0.07	0.078	0.07	0.120	0.07	0.215	10.00	0.09	0.52	4	4.1
Polyethylene glycol 1540	0.02	0.006	0.02	0.012	0.02	0.028	0.03	0.047	0.03	0.094					
Polyethylene glycol 4000	0.02	0.004	0.02	0.008	0.02	0.020	0.02	0.033	0.02	0.067					
Polymyxin B sulfate			0.09	0.052[c]			0.06	0.10	0.04	0.12					
Polysorbate 80	0.02	0.005	0.02	0.010	0.02	0.020	0.02	0.032	0.02	0.055					
Polyvinyl alcohol (99% hydrol)	0.02	0.004	0.02	0.008	0.02	0.020	0.02	0.035	0.03	0.075					
Polyvinylpyrrolidone	0.01	0.003	0.01	0.006	0.01	0.010	0.01	0.017	0.01	0.035					
Potassium acetate	0.59	0.172	0.59	0.342							1.53	0.59	0.52	0	7.6
Potassium chlorate											1.88			0	6.9
Potassium chloride			0.76	0.439[c]							1.19	0.76	0.52	0	5.9
Potassium iodide			0.34	0.196[c]							2.59	0.34	0.52	0	7.0
Potassium nitrate			0.56	0.324[c]							1.62	0.56		0	5.9
Potassium phosphate			0.46	0.27							2.08	0.43	0.52	0	8.4
Potassium phosphate, monobasic			0.44	0.25							2.18	0.41	0.52	0	4.4
Potassium sulfate			0.44								2.11	0.43		0	6.6
Pralidoxime chloride	0.32	0.092	0.32	0.183	0.32	0.364					2.87	0.32	0.52	0	4.6
Prilocaine HCl	0.22	0.062	0.22	0.125	0.22	0.250	0.22	0.375			4.18	0.22	0.52	45	4.6
Procainamide HCl			0.22	0.13			0.19	0.33	0.17	0.49					
Procaine HCl			0.21	0.122[c]			0.19	0.33	0.18		5.05	0.18	0.52	91	5.6
Prochlorperazine edisylate	0.08	0.020	0.06	0.033	0.05	0.048	0.03	0.056	0.02	0.065					
Promazine HCl	0.18	0.050	0.13	0.077	0.09	0.102	0.07	0.112	0.05	0.137					
Proparacaine HCl	0.16	0.044	0.15	0.086	0.15	0.169	0.14	0.247	0.13	0.380	7.46	0.12	0.52		
Propiomazine HCl	0.18	0.050	0.15	0.084	0.12	0.133	0.10	0.165	0.08	0.215					
Propoxycaine HCl											6.40			16	5.3
Propylene glycol											2.00			100	5.5
Pyrathiazine HCl	0.22	0.065	0.17	0.095	0.11	0.123	0.08	0.140	0.06	0.170					
Pyridostigmine bromide	0.22	0.062	0.22	0.125	0.22	0.250	0.22	0.377			4.13	0.22	0.52	0	7.2
Pyridoxine HCl											3.05			31*	3.2
Quinacrine methanesulfonate				0.06[c]											
Quinine bisulfate			0.09	0.05			0.09	0.16							
Quinine dihydrochloride			0.23	0.130[c]			0.19	0.33	0.18		5.07	0.18	0.52	trace*	2.5
Quinine hydrochloride			0.14	0.077[c]			0.11	0.19							
Quinine and urea HCl			0.23	0.13			0.21	0.36			4.5	0.20	0.52	64	2.9

Appendix A—Continued

	0.5%		1%		2%		3%		5%		Isoosmotic concentration[a]				
	E	D	E	D	E	D	E	D	E	D	%	E	D	H	pH
Resorcinol				0.161[c]							3.30			96	5.0
Rolitetracycline	0.11	0.032	0.11	0.064	0.10	0.113	0.09	0.158	0.07	0.204					
Rose Bengal	0.08	0.020	0.07	0.040	0.07	0.083	0.07	0.124	0.07	0.198	14.9	0.06	0.52		
Rose Bengal B	0.08	0.022	0.08	0.044	0.08	0.087	0.08	0.131	0.08	0.218					
Scopolamine HBr			0.12	0.07			0.12	0.21	0.12	0.35	7.85	0.11	0.52	8	4.8
Scopolamine methylnitrate			0.16				0.14		0.13		6.95	0.13		0	6.0
Secobarbital sodium			0.24	0.14			0.23	0.40			3.9	0.23	0.52	trace	9.8
Silver nitrate			0.33	0.190[c]							2.74	0.33	0.52	0*	5.0
Silver protein, mild			0.17	0.10			0.17	0.29	0.16	0.46	5.51	0.16	0.52	0	9.0
Silver protein, strong				0.06[d]											
Sodium acetate			0.46	0.267							2.0	0.45	0.52		
Sodium acetazolamide	0.24	0.068	0.23	0.135	0.23	0.271	0.23	0.406			3.86	0.23	0.52		
Sodium aminosalicylate				0.170[c]							3.27			0	7.3
Sodium ampicillin	0.16	0.045	0.16	0.090	0.16	0.181	0.16	0.072	0.16	0.451	5.78	0.16	0.52	0	8.5
Sodium ascorbate											3.00			0	6.9
Sodium benzoate			0.40	0.230[c]							2.25	0.40	0.52	0	7.5
Sodium bicarbonate			0.65	0.375							1.30	0.65	0.52	0	8.3
Sodium biphosphate (H_2O)			0.40	0.23							2.45	0.37	0.52	0	4.1
Sodium biphosphate (2 H_2O)			0.36								2.77	0.32		0	4.0
Sodium bismuth thioglycollate	0.20	0.055	0.19	0.107	0.18	0.208	0.18	0.303	0.17	0.493	5.29			0	8.3
Sodium bisulfite			0.61	0.35							1.5	0.61	0.52	0*	3.0
Sodium borate			0.42	0.241[c]							2.6	0.35	0.52	0	9.2
Sodium bromide											1.60			0	6.1
Sodium cacodylate			0.32				0.28				3.3	0.27		0	8.0
Sodium carbonate, monohydrated			0.60	0.346							1.56	0.58	0.52	100	11.1
Sodium cephalothin	0.18	0.050	0.17	0.095	0.16	0.179	0.15	0.259	0.14	0.400	6.80	0.13	0.52	partial	8.5
Sodium chloride			1.00	0.576[c]			1.00	1.73	1.00	2.88	0.9	1.00	0.52	0	6.7
Sodium citrate			0.31	0.178[c]			0.30	0.52			3.02	0.30		0	7.8
Sodium colistimethate	0.16	0.045	0.15	0.087	0.14	0.161	0.14	0.235	0.13	0.383	6.85	0.13	0.52	0	8.4
Sodium hypophosphite											1.60			0	7.3
Sodium iodide			0.39	0.222[c]							2.37	0.38	0.52	0	6.9
Sodium iodohippurate											5.92			0	7.3
Sodium lactate											1.72			0	6.5
Sodium lauryl sulfate	0.10	0.029	0.08	0.046	0.07	0.068	0.05	0.086							
Sodium mercaptomerin											5.30			0	8.4
Sodium metabisulfite			0.67	0.386[c]							1.38	0.65	0.52	5*	4.5
Sodium methicillin	0.18	0.050	0.18	0.099	0.17	0.192	0.16	0.281	0.15	0.445	6.00	0.15	0.52	0	5.8
Sodium nafcillin	0.14	0.039	0.14	0.078	0.14	0.156	0.13	0.219	0.10	0.285					
Sodium nitrate			0.68								1.36	0.66		0	6.0
Sodium nitrite			0.84	0.460[c]							1.08	0.83		0*	8.5
Sodium oxacillin	0.18	0.050	0.17	0.095	0.16	0.177	0.15	0.257	0.14	0.408	6.64	0.14	0.52	0	6.0
Sodium phenylbutazone	0.19	0.054	0.18	0.104	0.17	0.202	0.17	0.298	0.17	0.488	5.34	0.17	0.52		
Sodium phosphate			0.29	0.168			0.27	0.47			3.33	0.27	0.52	0	9.2
Sodium phosphate, dibasic (2 H_2O)			0.42	0.24							2.23	0.40	0.52	0	9.2
Sodium phosphate, dibasic (12 H_2O)			0.22				0.21				4.45	0.20		0	9.2
Sodium propionate			0.61	0.35							1.47	0.61	0.52	0	7.8
Sodium salicylate			0.36	0.210[c]							2.53	0.36	0.52	0	6.7
Sodium succinate	0.32	0.092	0.32	0.184	0.31	0.361					2.90	0.31	0.52	0	8.5
Sodium sulfate, anhydrous			0.58	0.34							1.61	0.56	0.52	0	6.2
Sodium sulfite, exsiccated			0.65	0.38							1.45			0	9.6
Sodium sulfobromophthalein	0.07	0.019	0.06	0.034	0.05	0.060	0.05	0.084	0.04	0.123					
Sodium tartrate	0.33	0.098	0.33	0.193	0.33	0.385					2.72	0.33	0.52	0	7.3
Sodium thiosulfate			0.31	0.181[c]							2.98	0.30	0.52	0	7.4
Sodium warfarin	0.18	0.049	0.17	0.095	0.16	0.181	0.15	0.264	0.15	0.430	6.10	0.15	0.52	0	8.1
Sorbitol (½ H_2O)											5.48			0	5.9
Sparteine sulfate	0.10	0.030	0.10	0.056	0.10	0.111	0.10	0.167	0.10	0.277	9.46	0.10	0.52	19*	3.5
Spectinomycin HCl	0.16	0.045	0.16	0.092	0.16	0.185	0.16	0.280	0.16	0.466	5.66	0.16	0.52	3	4.4
Streptomycin HCl			0.17	0.10[c]			0.16	0.16							
Streptomycin sulfate			0.07	0.036[c]			0.06	0.10	0.06	0.17					

Appendix A—Continued

	0.5%		1%		2%		3%		5%		Isoosmotic concentration[a]				
	E	D	E	D	E	D	E	D	E	D	%	E	D	H	pH
Sucrose			0.08	0.047[b]			0.09	0.16	0.09	0.26	9.25	0.10	0.52	0	6.4
Sulfacetamide sodium			0.23	0.132[c]			0.23	0.40			3.85	0.23	0.52	0	8.7
Sulfadiazine sodium			0.24	0.14			0.24	0.38			4.24	0.21	0.52	0	9.5
Sulfamerazine sodium			0.23	0.13			0.21	0.36			4.53	0.20	0.52	0	9.8
Sulfapyridine sodium			0.23	0.13			0.21	0.36			4.55	0.20	0.52	5	10.4
Sulfathiazole sodium			0.22	0.13			0.20	0.35			4.82	0.19	0.52	0	9.9
Tartaric acid				0.143[c]							3.90			75*	1.7
Tetracaine HCl			0.18	0.109[c]			0.15	0.26	0.12	0.35					
Tetracycline HCl			0.14	0.081[c]			0.10								
Tetrahydrozoline HCl											4.10			60*	6.7
Theophylline				0.02[b]											
Theophylline sodium glycinate											2.94			0	8.9
Thiamine HCl				0.139[c]							4.24			87*	3.0
Thiethylperazine maleate	0.10	0.030	0.09	0.050	0.08	0.089	0.07	0.119	0.05	0.153					
Thiopental sodium				0.155[c]							3.50			74	10.3
Thiopropazate diHCl	0.20	0.053	0.16	0.090	0.12	0.137	0.10	0.170	0.08	0.222					
Thioridazine HCl	0.06	0.015	0.05	0.025	0.04	0.042	0.03	0.055	0.03	0.075					
Thiotepa	0.16	0.045	0.16	0.090	0.16	0.182	0.16	0.278	0.16	0.460	5.67	0.16	0.52	10*	8.2
Tridihexethyl chloride	0.16	0.047	0.16	0.096	0.16	0.191	0.16	0.280	0.16	0.463	5.62	0.16	0.52	97	5.4
Triethanolamine	0.20	0.058	0.21	0.121	0.22	0.252	0.22	0.383			4.05	0.22	0.52	100	10.7
Trifluoperazine 2HCl	0.18	0.052	0.18	0.100	0.13	0.144									
Triflupromazine HCl	0.10	0.031	0.09	0.051	0.05	0.061	0.04	0.073	0.03	0.092					
Trimeprazine tartrate	0.10	0.023	0.06	0.035	0.04	0.045	0.03	0.052	0.02	0.061					
Trimethadione	0.23	0.069	0.23	0.133	0.22	0.257	0.22	0.378			4.22	0.21	0.52	100	6.0
Trimethobenzamide HCl	0.12	0.033	0.10	0.062	0.10	0.108	0.09	0.153	0.08	0.232					
Tripelennamine HCl				0.13[d]							5.50			100	6.3
Tromethamine	0.26	0.074	0.26	0.150	0.26	0.300	0.26	0.450			3.45	0.26	0.52	0	10.2
Tropicamide	0.10	0.030	0.09	0.050											
Trypan blue	0.26	0.075	0.26	0.150											
Tryparsamide				0.11[c]											
Tubocurarine chloride				0.076[c]											
Urea			0.59	0.34							1.63	0.55	0.52	100	6.6
Urethan				0.18[b]							2.93			100	6.3
Uridine	0.12	0.035	0.12	0.069	0.12	0.138	0.12	0.208	0.12	0.333	8.18	0.11	0.52	0*	6.1
Valethamate bromide	0.16	0.044	0.15	0.085	0.15	0.168	0.14	0.238	0.11	0.324					
Vancomycin sulfate	0.06	0.015	0.05	0.028	0.04	0.049	0.04	0.066	0.04	0.098					
Viomycin sulfate			0.08	0.05			0.07	0.12	0.07	0.20					
Xylometazoline HCl	0.22	0.065	0.21	0.121	0.20	0.232	0.20	0.342			4.68	0.19	0.52	88	5.0
Zinc phenolsulfonate											5.40			0*	5.4
Zinc sulfate			0.15	0.086[c]			0.13	0.23	0.12	0.35	7.65	0.12	0.52		

[a] The unmarked values were taken from Hammarlund *et al.*[14,17] and Sapp *et al.*[18]
[b] Adapted from Lund *et al.*[19]
[c] Adapted from BPC.[16]
[d] Obtained from several sources.
[e] *E:* sodium chloride equivalents. *D:* freezing-point depression, °C. *H:* hemolysis, %, at the concentration which is isoosmotic with 0.9% NaCl, based on freezing-point determination or equivalent (est. *pH:* approximate pH of solution studied for hemolytic action. *: change in appearance of erythrocytes and/or solution.[18,20] †: pH determined after addition of blood.

Appendix B—Volumes of Water for Isotonicity[21,a,b]

Drug (0.3 g)	Water needed for isotonicity, mL	Drug (0.3 g)	Water needed for isotonicity, mL	Drug (0.3 g)	Water needed for isotonicity, mL
Alcohol	21.7	Boric acid	16.7	Ephedrine sulfate	7.7
Ammonium chloride	37.3	Butacaine sulfate	6.7	Epinephrine bitartrate	6.0
Amobarbital sodium	8.3	Caffeine and sodium benzoate	8.7	Epinephrine hydrochloride	9.7
Amphetamine phosphate	11.3	Calcium chloride	17.0	Ethylmorphine hydrochloride	5.3
Amphetamine sulfate	7.3	Calcium chloride (6 H_2O)	11.7	Fluorescein sodium	10.3
Antipyrine	5.7	Chlorobutanol (hydrated)	8.0	Glycerin	11.7
Apomorphine hydrochloride	4.7	Chlortetracycline sulfate	4.3	Holocaine hydrochloride	6.7
Ascorbic acid	6.0	Cocaine hydrochloride	5.3	Homatropine hydrobromide	5.7
Atropine methylbromide	4.7	Cupric sulfate	6.0	Homatropine methylbromide	6.3
Atropine sulfate	4.3	Dextrose, anhydrous	6.0	Hyoscyamine sulfate	4.7
Bacitracin	1.7	Dibucaine hydrochloride	4.3	Neomycin sulfate	3.7
Barbital sodium	10.0	Dihydrostreptomycin sulfate	2.0	Oxytetracycline hydrochloride	4.3
Bismuth potassium tartrate	3.0	Ephedrine hydrochloride	10.0	Penicillin G, potassium	6.0

Appendix B—Continued

Drug (0.3 g)	Water needed for isotonicity, ml.	Drug (0.3 g)	Water needed for isotonicity, ml.	Drug (0.3 g)	Water needed for isotonicity, ml.
Penicillin G, sodium	6.0	Scopolamine methylnitrate	5.3	Sodium propionate	20.3
Pentobarbital sodium	8.3	Secobarbital sodium	8.0	Sodium sulfite, exsiccated	21.7
Phenobarbital sodium	8.0	Silver nitrate	11.0	Sodium thiosulfate	10.3
Physostigmine salicylate	5.3	Silver protein, mild	5.7	Streptomycin sulfate	2.3
Pilocarpine hydrochloride	8.0	Sodium acetate	15.3	Sulfacetamide sodium	7.7
Pilocarpine nitrate	7.7	Sodium bicarbonate	21.7	Sulfadiazine sodium	8.0
Piperocaine hydrochloride	7.0	Sodium biphosphate, anhydrous	15.3	Sulfamerazine sodium	7.7
Polymyxin B sulfate	3.0			Sulfapyridine sodium	7.7
Potassium chloride	25.3	Sodium biphosphate	13.3	Sulfathiazole sodium	7.3
Potassium nitrate	18.7	Sodium bisulfite	20.3	Tetracaine hydrochloride	6.0
Potassium phosphate, monobasic	14.7	Sodium borate	14.0	Tetracycline hydrochloride	4.7
		Sodium iodide	13.0	Viomycin sulfate	2.7
Procainamide hydrochloride	7.3	Sodium metabisulfite	22.3	Zinc chloride	20.3
Procaine hydrochloride	7.0	Sodium nitrate	22.7	Zinc sulfate	5.0
Scopolamine hydrobromide	4.0	Sodium phosphate	9.7		

[a] Table of "Isotonic Solution Values" showing volumes in mL of solution that can be prepared by dissolving 300 mg of the specified drug in sterile water. The addition of an isotonic vehicle (commonly referred to as diluting solution) to make 30 mL yields a 1% solution. Solutions prepared as directed above are isosmotic with 0.9% sodium chloride solution but may not be isotonic with blood (see Appendix A for hemolysis data).

[b] To calculate V values for drugs which do not appear in Appendix B, but are listed in Appendix A, simply take the appropriate sodium chloride equivalent (E) and multiply by 0.3g to convert that quantity of drug to the equivalent weight of sodium chloride and divide by 0.009g (the weight of sodium chloride which will render 1 mL of water isotonic).

Example—Calculate the V value for anileridine HCl (Appendix A defines E = 0.18).

$$\frac{0.3 \times 0.18}{0.009} = 6 \text{ mL water for each 0.3 g drug}$$

CHAPTER 82

Quality Assurance and Control

Clyde R Erskine, Jr
Vice President
Corporate Quality Audits and Services
SmithKline Beckman Corp
One Franklin Plaza, PO Box 7929
Philadelphia, PA 19101

The Pharmaceutical Industry continues as a vital segment of the health-care cycle in conducting research and manufacturing products which are life-maintaining and life-restoring. The last decade has seen an evolution in the concepts relating to the Quality Assurance and Control of these products.

The changes brought about in assuring the safety and therapeutic efficacy of drug products have resulted from a number of factors which are either internal or external to the industry. Internally are the self-designed guidelines the industry has imposed on itself, exemplified by a document prepared in 1967 by the Pharmaceutical Manufacturers Association (PMA) titled "General Principles of Total Quality Control in the Drug Industry." This PMA document became the basis for later regulatory Guidelines prepared by the Food and Drug Administration (FDA) titled "Current Good Manufacturing Practice in Manufacture, Processing, Packing or Holding of Human and Veterinary Drugs." These Current Good Manufacturing Practices (CGMPs) have become the primary external guidelines used by industry and the FDA in the control and inspection of manufacturing facilities.

Quality Control and Assurance Organization

Although the terms Quality Control and Quality Assurance often are used interchangeably, depending on the structure of a specific company, there is a continuing trend to separate and define their functional responsibilities.

Quality Control can be defined broadly as the day-to-day control of quality within a company, a department staffed with scientists and technicians responsible for the acceptance or rejection of incoming raw materials and packaging components, for the myriad of in-process tests and inspections, to assure that systems are being controlled and monitored and, finally, for the approval or rejection of completed dosage forms.

Quality Control, therefore, includes not only the analytical testing of the finished product, but also the assessment of all operations beginning with the receipt of raw materials and continuing throughout the production and packaging operations, finished product testing, documentation, surveillance and distribution.

Quality Assurance may be defined as the responsibility of an organization to determine that systems, facilities and written procedures both are adequate and followed in order to assure that products are controlled and will meet, in the final dosage form, all the applicable specifications. Quality Assurance naturally then becomes an oversight function, often auditing operations to determine that procedures and systems are suitable and, if not, to recommend the required changes. Higher management looks toward the Quality Assurance unit in order to develop some level of "comfort" as to how well they are meeting company standards and applicable government regulations.

Total Quality Control

The high quality of pharmaceutical products results from meticulous adherence to written procedures in carrying out all operations, beginning with research. It is at this early point that the quality begins to be designed into a product. Raw materials must be characterized and purchased from reputable suppliers so that uniform, stable products will result when these materials are incorporated into the finished dosage form. Facilities must be designed, systems installed and the proper equipment selected so that the potential for cross contamination of one product by another is eliminated, that material flow and personnel movements are planned to reduce the potential for product mix-ups and that the air and water, which is being provided to production, is adequate in amount and quality for the particular operations being performed.

Production personnel must be trained properly to perform their jobs, and the directions they follow must be written, approved by responsible individuals and adhered to strictly.

Shipping departments are responsible for seeing that the products are protected from adverse handling and environmental conditions while in transit to distribution points and customers.

Quality Control is ever-present, overseeing each of these operations and giving the final release approval for distribution only after assessing and being satisfied that each step in this process has been completed correctly.

These principles were highlighted in that original PMA document from which the following excerpt is taken:

"The quality of a product is its degree of possession of those characteristics designed and manufactured into it which contribute to the performance of an intended function when the product is used as directed. The quality of medicinal and related products is the sum of all factors which contribute directly or indirectly to the safety, effectiveness and acceptability of the product. Quality must be built into the product during research, development and production.

"Total control of quality as it applies to the drug industry is the organized effort within an entire establishment to design, produce, maintain and assure the specified quality in each unit of product distributed. The effort should not only establish specifications for product acceptance but should provide procedures and methods for achieving conformance with such specifications.

"The large variety of substances used in this industry, the complexity of its products and the various types of company organization make it impossible to design in detail a single universally applicable system for the total control of quality.

"The ultimate objective of a program for the total control of quality in a drug company is the attainment of perfection in meeting specifications for a product of high quality. It is a program designed to assure the professional user or ultimate consumer that every lot of a product conforms to specifications and that each dose distributed will fulfill the representations made in the labeling and will meet all legal requirements and such additional standards as the management of a firm may adopt.

"Total control of quality is a plantwide activity and represents the aggregate responsibility of all segments of a company. The responsibility for auditing the control system and for evaluating product quality is that of a specific group referred to in this statement as Quality Control. The head of Quality Control should have the authority to release satis-

factory lots of products, to reject unsuitable lots and to recommend the recall from distribution of any lots subsequently found to be unsuitable. He should be responsible to a level of management which enables him to exercise independent judgment. His responsibility and authority should be clearly defined by management."

It readily becomes apparent that quality must be built into a product and that it cannot be inspected or tested into a product. Quality results from teamwork, an association which is becoming increasingly important as the industry advances in new technologies which themselves are becoming more complex and demanding.

Quality Control and Assurance Functions

The head of Quality Control, who is ultimately responsible for decisions relating to the acceptability of finished product, should report to someone other than the person directly responsible for producing the product. Often in current organizational structures, the persons in charge of both quality control and production will report to some higher level of authority. This may be the same or different individuals, but it does allow for the independent operation of both functions without direct conflict arising when reaching the ultimate decision on the acceptability of product. The Quality Control function in an organization normally consists of at least two primary units, analytical control and inspection control.

Analytical Control

The Analytical Control Laboratory is responsible for testing and approving raw materials, work in-process and finished product. The laboratory must be staffed with persons who are trained both academically and by experience to perform the often complex analyses required to evaluate the acceptability of a product. Proper personnel is not the only necessity in the laboratory. Equipment also is required which will allow timely and accurate analysis. This equipment continues to become more sophisticated, providing more information about compounds than previously known and has led to a level of accuracy and detectability heretofore unknown.

Detailed specifications also must be available, as well as the test methods against which the products are measured. The specifications include the criteria against which the product will be evaluated and the limits for acceptance or rejection for each critical parameter.

The testing and acceptance of only high-quality raw materials is essential in the preparation of products. Part of this acceptance is to purchase raw materials only from known, reputable suppliers. In order to assure this condition, it is essential that Quality Control be part of a preapproval program of all potential suppliers. This approval always includes testing the material and in many cases necessitates an inspection of the supplier's facility to determine its suitability and degree of compliance with GMP's. At various critical in-process production or intermediate steps it may be necessary to sample and test the materials against criteria previously established for that particular step in the process.

Often, in-process alert or action levels will be identified at the critical operational steps as a means of process control. These alert or action levels are limits or specifications which are more restrictive than the final acceptance limits, but serve as in-process controls by giving early warnings of conditions which could lead to an out-of-control situation and allow timely corrective action to be taken before this occurs. Thus, materials reaching the alert or action-level criteria are acceptable, since they have not exceeded a rejection or unacceptable level.

In-process critical testing will vary depending on the dosage form being manufactured. Sterile parenteral products probably receive the most critical in-process control and testing in order to insure a finished product which is sterile and free of microbial contamination and particulate matter. With sterile products the end product sterility testing cannot be relied upon to insure that each and every container in a lot of an injectable product is sterile and dependence is placed on in-process controls. These in-process controls must have been developed following a prescribed protocol which defines operating conditions and parameters. Only after a series of successful production runs, using the prescribed parameters, can a process be judged to have been validated. Validation of processes is a critical step in the quality assurance of both sterile and nonsterile products. Validation may be defined as "assurance that production processes are controlled in such a manner that they will perform routinely in the manner in which they are purported to."

Testing of the completed lot of a dosage form, in order to measure its conformance with predetermined specifications and appropriate acceptance criteria, always is desirable before releasing the lot for shipment. However, the use of a properly validated manufacturing process is more critical to the quality of a product. End product testing suffers due to the normal variations that arise in the statistical sampling of a lot in assuring that a sample is homogeneous and representative.

Validation of processes and systems gives increased assurance of finished product lot quality and is leading the way toward reducing or eliminating the reliance on end product testing. The parametric release of finished product is beginning, based on control of the critical elements of a validated process.

Tests and specifications may be found in several sources. The *United States Pharmacopeia/National Formulary* (USP) is published on a 5-yr cycle program by the United States Pharmacopeial Convention. The standards established by and published in the USP have been recognized as being official by the Congress of the United States and are recognized in the Federal Food, Drug and Cosmetic Act. These standards are prepared and reviewed so that through regular revision, entirely or in part, they remain current. The reviewing body known as the Committee of Revision represents medical, academe, industrial and other scientific experts. The primary purpose of the Committee is "to provide authoritative standards for materials and substances and their preparations that are used in the healing arts." They establish titles, definitions, descriptions and standards for the identity, quality, strength, purity and, where practical, methods for their examination.

In addition to the procedures defined in the USP, companies will prepare their own test specifications when the products are not "official" (eg, not listed in the USP). These tests and specifications form a necessary part of the Control Sections of New Drug Applications (NDAs) submitted to the Federal Government and which, following careful review by the FDA, may be approved. Finally, there are test procedures for unofficial products and for those not requiring the submission of an NDA. Companies in these cases prepare their own in-house test procedures for controlling the products they produce.

Inspection Control

Many responsibilities assumed by Quality Control are ancillary to the analytical testing. These include the sampling and inspection of incoming raw materials, packaging and labeling components; the physical inspection of product at various intermediate stages; packaging line inspection and the control of shipping inventory within the distribution

cycle. Depending on the organizational structure, additional or different responsibilities will be assigned to this unit.

Documentation

During the course of producing a pharmaceutical product, numerous documents and records are generated. Each batch is assigned a specific code or lot number. All documentation relating to a specific code is referred to as a "batch record," which will include data on each significant phase of production, control and distribution. The batch record provides a historical blueprint of every step, beginning with the receipt of chemical raw materials and packaging components and continuing through each in-process stage. Recording charts or computer printouts of significant operations such as autoclaving, drying, air-particulate monitoring, lyophylizing, etc, all become part of this batch history. After the batch has been completed, including final analytical and physical testing, one additional step should be completed prior to approving the lot for distribution. All documents and records relating to the specific batch are given a final review. Each required document in the batch record must be checked for completeness and accuracy. Any discrepancy must be investigated immediately and answered. Only after this review has been completed satisfactorily may the batch be released for distribution.

When the batch has been released, accurate shipping records must be maintained showing the batch distribution. With these records it is then possible to trace the batch to the market place which will facilitate, if the need arose, recalling the product (batch) from the market place.

Quality Assurance

Total control of quality not only requires the assignments described above, but should include a monitoring or audit function as well. The responsibility for this function is normally separate from both the production and control operations, thus allowing an independent oversight of all operations. Although the function may be separated, the audit responsibilities are often shared by a team representing both the production and control disciplines. It is the duty of this individual (or team) through review and inspection to assure that written procedures and policies are available for each significant production and control operation. Normally, standard operating procedures (SOPs) are developed which, when followed by properly trained operators, will help to assure the quality and integrity of the product. Thus, the QA review function not only determines that the procedures are current and correct, but that they are being followed. Combining a review of SOPs with an audit of facilities and operations following the applicable GMP regulations will give a company an "inside" report on its level of compliance and will allow necessary changes and/or corrections to be made prior to either causing a product failure or being observed during an inspection by an FDA investigator.

Production is responsible for following prescribed procedures to produce acceptable products. The system of total quality management becomes the joint responsibility of quality control and quality assurance.

Quality depends, to a major degree, upon the employees engaged in the production operations. They are responsible for following the prescribed procedures and, along with their training and experience, are able to produce uniformly acceptable products. GMPs properly organized and followed afford a mechanism for preventing human error, the potential for which is especially great in this industry.

New Advances

Statistics and trend analysis are tools already used by the pharmaceutical industry in determining the proper sample size required for testing, for measuring the uniformity of solid dosage forms and for plotting trends of significant factors in order to correct out-of-control situations before unacceptable product results.

New management concepts are being tested, directed toward a reduction of raw material inventories and packaging components. The term "Just-In-Time" refers to ordering and receiving materials when they are required for production rather than to maintaining extensive inventories. This places an additional burden on Quality Control and suppliers to assure the acceptability of materials when received. Certification and qualification of suppliers is an expanding responsibility of control personnel.

Electronic data processing has become another useful tool for assessing process and test parameters and for analyzing the data collected during production. The control of many operations by computers and microprocessors is providing the capability for producing products of further improved uniform quality. These systems have challenged the older ones, resulting in new approaches to in-process controls, collection and analysis of data and the entire system of quality control.

Robotics is finding various applications in pharmaceutical production, packaging and laboratory operations. Filling of product into containers, cartoning, palletizing and other material handling tasks as well as laboratory testing and sample preparation are either in use or being investigated. The uniformity of procedures, costs, nonfatigue factors and flexibility are all advantages. Probably the most practical use of robots is in sterile processing where their nonevasiveness allows aseptic production and testing by removing a primary source of contamination, the human worker. These all present the potential for improved control systems while bringing new challenges to the QC professional.

Environmental Control

Along with the many other advances in the total control of quality is the growing recognition that the environment and the systems used for its control can have a significant effect on the finished product quality. It is well-recognized that parenteral or sterile ophthalmic products must be produced in a manner which will insure their sterility; therefore, control of the areas in which they are manufactured is essential.

Microbiological monitoring of air and water to control the level of particulate and microbial matter in these production areas is necessary. Several levels of "clean" areas are described in Federal Government Standard 209C, "Clean Room and Work Station Requirements, Controlled Environment." The industry commonly uses the specifications described which classify air cleanliness based on the number of particles (of a given size) per cubic foot of air. Generally, conditions listed as "Class 100" are maintained in areas where parenteral products are filled into sterile containers. Class 100 is defined as an area which can be maintained at less than 100 particles per cubic foot of air 0.5 μm and larger.

Another essential control procedure is microbiological monitoring of the environment in which nonsterile products are manufactured. The objective of this monitoring is to first determine particulate and microbial levels within an area to assure that they are reasonable. If found to be excessive, steps must be taken to bring the levels to within acceptable limits. Once this base has been developed, regular monitoring will indicate if operations are continuing under acceptable limits. If not, immediate corrective action should be taken.

The monitoring and control of particulate and microbial matter will further assure the final quality and stability of the product because the environment has been controlled and the product has not been challenged by an unacceptable level of particulate generated by an out-of-control situation.

Good Manufacturing Practice Regulations

In June, 1963, the FDA first issued regulations describing the current good manufacturing practice to be followed in the manufacture, packaging and holding of finished pharmaceuticals. The regulations underwent significant revision and updating in 1978 and became official in March, 1979. These regulations present the minimum requirements to be met by industry when manufacturing, processing, packaging and holding of human and veterinary drugs. Under the Federal Food, Drug and Cosmetic Act, a drug is deemed to be adulterated unless the methods used in its manufacture, processing, packing and holding, as well as the facilities and controls used, conform to current good manufacturing practice so that the drug meets the safety requirements of the Act and has the identity and strength to meet the quality and purity characteristics that it is represented to have. In the preamble to the regulations, the FDA Commissioner answers the comments received from interested persons who responded when the proposed rules were first issued. The preamble provides interesting background information as to why specific sections of the regulations were believed to be necessary and their interpretation.

In July, 1978, the FDA issued regulations establishing similar Good Manufacturing Practices (GMPs) for the Manufacture, Packing, Storage and Installation of Medical Devices. These were published following an amendment to the Food, Drug and Cosmetic Act of 1976, which provided the FDA with the authority to prescribe regulations pertaining to medical devices. In December, 1978, regulations concerning Good Laboratory Practices (GLP) for the control and conducting of clinical studies were issued and for the first time came under FDA inspectional authority.

The FDA proposed, in June 1978, regulations covering the GMPs relating to the manufacture and control of large-volume parenteral products. These regulations, although never officially issued, have become the guideline used by the industry and FDA in the manufacture, control and inspection of large-volume parenteral production. Due to the similarity of the controls required for the production of small-volume parenterals, the guidelines also have been used to assess the adequacy of the manufacture and controls used with these products.

A number of other "guidelines" or "concept" papers have been prepared by various organizations within the industry itself, such as the Pharmaceutical Manufacturers Association and the Parenteral Drug Association. A partial listing is provided at the end of this section.

The current GMP regulations should be read and understood thoroughly by those involved in or interested in pursuing quality control or quality assurance responsibilities. The scope of the present regulations is given in the following outline, along with a brief interpretation of each section.

Bibliography

Human and veterinary drugs—current good manufacturing practice in manufacture, processing, packing or holding. *21 CFR 211*: 1982.

General Principles of Total Quality Control in the Drug Industry, PMA, Washington DC, June 1967.

Clean room and work station requirements, controlled environment. *Fed Std 209C* (Proposed), GSA, Washington DC, Sept. 1986.

Good manufacturing practice for medical devices, general. *21 CFR 820*: Apr 1982.

Good laboratory practice for non-clinical laboratory studies. *21 CFR 58*: Apr 1982.

Human drugs—current good manufacturing practice in manufacture, processing, packing or holding large volume parenterals. *Fed Reg 41(106)*: June 1, 1976.

USP XXI/NF XVI: 1985.

Validation of Steam Sterilization Cycles (Tech Mono No 1), PDA, Philadelphia, 1978.

Validation of Aseptic Filling for Solution Drug Products (Tech Mono No 2), PDA, Philadelphia, 1980.

Validation of Dry Heat Processes Used for Sterilization and Depyrogenation (Tech Mono No 3), PDA, Philadelphia, 1981.

Proceedings for the PMA Seminar Program on Validation of Solid Dosage Form Processes, PMA, Washington DC, May 1980.

Validation of Sterilization of Large Volume Parenterals—Current Concepts, PMA, Washington DC, Feb 1979.

Guideline on Sterile Drug Products Produced by Aseptic Processing, FDA, Rockville MD, June 1987.

Proceedings of the PMA Seminar Program on Concepts and Principles for the Validation of Computer Systems Used in the Manufacture and Control of Drug Products, PMA, Washington DC, Apr 1988.

PART 211—CURRENT GOOD MANUFACTURING PRACTICE IN MANUFACTURE, PROCESSING, PACKING OR HOLDING—HUMAN AND VETERINARY DRUGS

Subpart A—General Provisions		
211.3	Definitions	*Interpretation* The scope of the regulations are explained for human prescription and OTC drug products including biological products. Reference is made to Part 210.3 of the chapter which gives definitions for all significant terms used in the regulations.
Subpart B—Organization and Personnel		
211.22	Responsibilities of quality control unit	Highlighted here is the assignment to the quality control unit total responsibility for ensuring that adequate systems and procedures exist and are followed to assure product quality.
211.25	Personnel qualifications	Personnel, either supervisory or operational, must be qualified by training and experience to perform their assigned tasks.
211.28	Personnel responsibilities	The obligations of personnel engaged in the manufacture of drug products concerning their personal hygiene, clothing and medical status are defined.
211.34	Consultants	The qualifications of consultants must be approved by Quality Control.
Subpart C—Buildings and Facilities		
211.42	Design and construction features	Buildings and facilities can be considered acceptable only if they are suitable for their intended purpose and can be maintained. Construction concepts, such as air handling systems, lighting, eating facilities and plumbing systems including water, sewage and toilet facilities, are outlined.
211.44	Lighting	
211.46	Ventilation, air filtration, air heating and cooling	
211.48	Plumbing	
211.50	Sewage and refuse	
211.52	Washing and toilet facilities	
211.56	Sanitation	
211.58	Maintenance	

Subpart D—Equipment		
211.63	Equipment design, size and location	Equipment must be designed, constructed, of adequate size, suitably located and able to be maintained in order to be considered suitable for its intended use.
211.65	Equipment construction	
211.67	Equipment cleaning and maintenance	
211.68	Automatic, mechanical and electronic equipment	Reference is made to the use of automatic equipment, data processors and computers highlighting the need to verify output versus input and for proper calibration of recorders, counters and other electrical or mechanical devices.
211.72	Filters	Special note is made that only filters are to be used which do not release fibers into products.
Subpart E—Control of Components and Drug Product Containers and Closures		
211.80	General requirements	Written procedures must be available which describe the receipt, identification, storage, handling, sampling, testing and approval or rejection of components (raw materials) and drug products.
211.82	Receipt and storage of untested components, drug product containers and closures	
211.84	Testing and approval or rejection of components, drug product containers and closures	Once approved or rejected, these materials must be so identified and stored. If approved, they must be inventoried in a manner to assure that the oldest approved stock is used first (FIFO). Materials which are subject to deterioration during storage should be retested at an appropriate time based on stability profiles.
211.86	Use of approved components, drug product containers and closures	
211.87	Retesting of approved components, drug product containers and closures	
211.89	Rejected components, drug product containers and closures	
211.94	Drug product containers and closures	Containers and closures (product contact materials) must be nonreactive with or additive to the product.
Subpart F—Production and Process Controls		
211.100	Written procedures; deviations	Written standard operating procedures (SOP's) for each production process and control procedure are necessary. Any deviation to a SOP must be investigated, recorded and approved prior to final product acceptance.
211.101	Charge-in of components	All products are to be formulated to provide not less than 100% of the required amount of active ingredient. Records are to be maintained of each component and the quantity which is incorporated into a batch.
211.103	Calculation of yield	
211.105	Equipment identification	
211.110	Sampling and testing of in-process materials and drug products	Significant in-process steps are to be identified and appropriate sampling, testing and approvals obtained before proceeding further in the production cycle. If required, time limitations will be placed on in-process steps.
211.111	Time limitations on production	
211.113	Control of microbiological contamination	Appropriate procedures are to be prepared for testing components, products and the environment in order to establish that a product is not microbiologically contaminated.
211.115	Reprocessing	Reprocessing of product is allowed providing there are written procedures covering the methods to be used and approved by quality control. Additional testing of the reprocessed batch may be required to assure conformity with specifications.
Subpart G—Packaging and Labeling Control		
211.122	Materials examination and usage criteria	Labeling & packaging materials are to be received, identified, stored, sampled and tested following detailed written procedures.
211.125	Labeling issuance	Special controls must be exercised over labeling to assure that only the correct labels are issued to packaging for a specific product and that the quantities used are reconciled with the quantity issued.
211.130	Packaging and labeling operations	
211.134	Drug product inspection	
211.137	Expiration dating	Following appropriate stability studies at prescribed temperature conditions, products on the market shall bear an expiration date to assure that they are used within their expected shelf life.
Subpart H—Holding and Distribution		
211.142	Warehousing procedures	Describes the requirements for warehousing and distribution of products and their holding under appropriate conditions of light, temperature and humidity.
211.150	Distribution procedures	
Subpart I—Laboratory Controls		
211.160	General requirements	Concerns written procedures in the form of specifications, standards, sampling plans and test procedures which are used in a laboratory for controlling components and finished drug products. Acceptance criteria for sampling and approval shall be adequate for support release of product to distribution.
211.165	Testing and release for distribution	
211.166	Stability testing	A stability testing program will be followed in order to assess the stability characteristics of drug products. The results of this testing shall be used in assigning appropriate storage conditions and expiration dates.
211.167	Special testing requirements	
211.170	Reserve samples	
211.173	Laboratory animals	Animals used in any testing shall be maintained and controlled in a manner suitable for use.
211.176	Penicillin contamination	Drug products cannot be marketed if, when tested by a prescribed procedure, found to contain any detectable levels of penicillin.
Subpart J—Records and Reports		
211.180	General requirements	Details the various records and documents which should be generated during the manufacture of drug products and which are to be available for review.
211.182	Equipment cleaning and use log	
211.184	Component, drug product container, closure and labeling records	
211.186	Master production and control records	A master production record must be prepared for each drug product, describing all aspects of its manufacture, packaging and control. Individual batch records are derived from this approved master.
211.188	Batch production and control records	
211.192	Production record review	
211.194	Laboratory records	

211.196 Distribution records	Distribution records include warehouse shipping logs, invoices, bills of lading and all documents associated with distribution. These records should provide all the information necessary to trace lot distribution in order to facilitate product retrieval if necessary.
211.198 Complaint files	Records of complaints received from consumers and professionals are to be maintained along with the report of their investigation and response.
Subpart K—Returned and Salvaged Drug Products	
211.204 Returned drug products	Records are to be maintained of drug products returned from distribution channels and the reason for their return. This data can be used as part of the total lot accountability, should the need arise, to trace its distribution and/or for its recall.
211.208 Drug product salvaging	Drug products that have been stored improperly are not to be salvaged.

CHAPTER 83

Solutions, Emulsions, Suspensions and Extracts

J G Nairn, PhD
Professor of Pharmacy
Faculty of Pharmacy
University of Toronto
Toronto, Canada M5S 1A1

The dosage forms described in this chapter may be prepared by dissolving the active ingredient(s) in an aqueous or nonaqueous solvent, by suspending the drug (if it is insoluble in pharmaceutically or therapeutically acceptable solvents) in an appropriate medium or by incorporating the medicinal agent into one of the two phases of an oil and water system. Such solutions, suspensions and emulsions are further defined in subsequent paragraphs but some, with similar properties, are considered elsewhere. These dosage forms are useful for a number of reasons. They can be formulated for different routes of administration: oral use, introduction into body cavities or applied externally. The dose easily can be adjusted by dilution, and the oral liquid form readily can be administered to children or people unable to swallow tablets or capsules. Extracts eliminate the need to isolate the drug in pure form, allow several ingredients to be administered from a single source (eg, pancreatic extract) and permit the preliminary study of drugs from natural sources. Occasionally, solutions of drugs such as potassium chloride are used to minimize adverse effects in the gastrointestinal tract.

The preparation of these dosage forms involves several considerations on the part of the pharmacist: purpose of the drug, internal or external use, concentration of the drug, selection of the liquid vehicle, physical and chemical stability of the drug, preservation of the preparation and use of appropriate excipients such as buffers, solubilizers, suspending agents, emulsifying agents, viscosity controlling agents, colors and flavors. The theory of many of these preparations is discussed in earlier chapters in Part 2, *Pharmaceutics*. Because of the complexity of some manufactured products, compounding may be carried out with the aid of linear programming models in order to obtain the optimal product. The appropriate chapters (see the index) should be consulted for information on the preparation and characteristics of those liquid preparations that are intended for ophthalmic or parenteral use.

Much has been written during the past decade about the biopharmaceutical properties of, in particular, the solid dosage forms. In assessing the bioavailability of drugs in tablets and capsules, many researchers first have studied the absorption of drugs administered in solution. Since drugs are absorbed in their dissolved state, frequently it is found that the absorption rate of oral dosage forms decreases in the following order: aqueous solution > aqueous suspension > tablet or capsule. The bioavailability of a medicament, for oral ingestion and absorption, should be such that eventually all of the drug is absorbed as it passes through the gastrointestinal tract, regardless of the dosage form. There are a number of reasons for formulating drugs in forms in which the drug is not in the molecular state. These are: improved stability, improved taste, low water solubility, palatability and ease of administration. It becomes apparent, then, that each dosage form will have advantages and disadvantages.

The pharmacist handles liquid preparations in one of three ways. He may dispense the product in its original container, buy the product in bulk and repackage it at the time a prescription is presented by the patient or compound the solution, suspension or emulsion in the dispensary. Compounding may involve nothing more than mixing marketed products in the manner indicated on the prescription or, in specific instances, may require the incorporation of active ingredients in a logical and pharmaceutically acceptable manner into the aqueous or nonaqueous solvents which will form the bulk of the product.

The pharmacist, in the first instance, depends on the pharmaceutical manufacturer to produce a product that is effective, elegant and stable when stored under reasonably adverse conditions. Most manufacturers attempt to guarantee efficacy by evaluating their products in a scientifically acceptable manner but, in some instances, such efficacy is relative. For example, cough mixtures marketed by two different manufacturers may contain the same active ingredients and it becomes difficult to assess the relative merits of the two products. In such instances the commercial advantage gained by one over the other may be based on product elegance. Thus, color, odor, taste, pourability and homogeneity are important pharmaceutical properties.

The stability of the active ingredient in the final product is of prime concern to the formulator. In general, drug substances are less stable in aqueous media than in the solid dosage form and it is important, therefore, to properly buffer, stabilize or preserve, in particular those solutions, suspensions and emulsions that contain water. Certain simple chemical reactions can occur in these products. These may involve an ingredient-ingredient interaction (which implies a poor formulation), a container-product interaction (which may alter product pH and thus, for pH-sensitive ingredients, be responsible for the subsequent formation of precipitates) or a direct reaction with water (ie, hydrolysis). The stability of pharmaceutical products is discussed in Chapter 81.

The more complicated reactions usually involve oxygen. Vitamins, essential oils and almost all fats and oils can be oxidized. Formulators usually use the word *autoxidation* when the ingredient(s) in the product react with oxygen but without drastic external interference. Such reactions first must be initiated by heat, light (including ultraviolet radiant energy), peroxides or other labile compounds or heavy metals such as copper or iron. This initiation step results in the formation of a free radical (R*) which then reacts with oxygen.

$$R^* + O_2 \rightarrow RO_2^* \text{ (peroxy radical)}$$

$$RO_2^* + RH \rightarrow ROOH + R^*$$

The free radical thus is regenerated and reacts with more oxygen. This propagation step is followed by the termination reactions.

$$RO_2^* + RO_2^* \rightarrow \text{inactive product}$$

$$RO_2^* + R^* \rightarrow \text{inactive product}$$

$$R^* + R^* \rightarrow \text{inactive product}$$

The effect of trace metals can be minimized by using citric acid or EDTA (ie, sequestering agents). Antioxidants, on the other hand, may retard or delay oxidation by reacting with the free radicals formed in the product. Examples of antioxidants are the propyl, octyl and dodecyl esters of gallic acid, butylated hydroxyanisole (BHA) and the tocopherols or vitamin E. For a more detailed approach to the prevention of oxidative deterioration in pharmaceuticals, the papers by Ostendorf[1] and Chalmers,[2] should be consulted. A description of many antioxidants is given in Chapter 66.

The problem of drug stability has been well-defined by pharmaceutical scientists but during the past few years a secondary and, in some respects, more serious problem has confronted the manufacturer of liquid preparations. Such pharmaceutically diverse products as baby lotions and milk of magnesia have been recalled from the market because of microbial contamination. In a survey of retail packages of liquid antacid preparations containing magnesium hydroxide, it was found that 30.5% of the finished bottles were contaminated with *Pseudomonas aeruginosa*. The aerobic plate count ranged from less than 100 to 9,300,000 organisms/g. Other examples could be cited but the range of microorganisms which can contaminate the liquid preparation includes the *Salmonella* sp, *E coli*, certain *Pseudomonas* sp, including *P aeruginosa*, and *Staphylococcus aureus*. Bruch[3] describes the types of microorganisms found in various products and attempts to evaluate the hazards associated with the use of nonsterile pharmaceuticals. Coates[4] in a series of papers describes various interactions which must be considered when preservatives are selected.

The USP recommends that certain classes of products be tested routinely for microbial contamination, eg, natural plant, animal and some mineral products, for freedom from *Salmonella* sp; oral solutions and suspensions, for freedom from *E coli*; articles applied topically, for freedom from *P aeruginosa* and *S aureus* and articles for rectal, urethral or vaginal administration, for total microbial count.

Products may become contaminated for a number of reasons.

The raw materials used in the manufacture of solutions, suspensions and emulsions are excellent growth media for bacteria. Water, in particular, must be handled with care but substances such as gums, dispersing agents, surfactants, sugars and flavors can be the carriers of bacteria which ultimately contaminate the product.

Equipment. Bacteria grow well in the nooks and crevices of pharmaceutical equipment (and in the simple equipment used in the dispensary). Such equipment should be cleaned thoroughly prior to use.

Environment and personnel can contribute to product contamination. Hands and hair are the most important carriers of contaminants. General cleanliness thus is vital. Head coverings must be used by those involved in the manufacturing process and face masks should be used by those individuals suffering from colds, coughs, hay fever and other allergic manifestations.

Packaging should be selected so that it will not contaminate the product and also will protect it from the environment.

The factors cited above relate to good manufacturing practice. However, the formulator can add a preservative to the product and decrease the probability of product contamination. If the product contains water, it almost is mandatory to include a preservative in the formulation. It must be stressed that this in no way replaces good in-plant control but merely provides further assurance that the product will retain its pharmaceutically acceptable characteristics until it is used by the patient.

The major criteria that should be considered in selecting a preservative: it should be effective against a wide spectrum of microorganisms, stable for its shelf life, nontoxic, nonsensitizing, compatible with the ingredients in the dosage form and relatively free of taste and odor.

Preservatives may be used alone or in combination to prevent the growth of microorganisms. Ethanol is a highly effective preservative. It is used at the 15% level in acidic media and at the 18% level in neutral or slightly alkaline media. Isopropyl alcohol is a fairly effective agent but it can be used only in topical preparations. Propylene glycol, a dihydric alcohol, has germicidal activity similar to that of ethanol. It normally is used in a 10% concentration.

A 0.5% solution of phenol is a good preservative but it is toxic, has its own characteristic odor and reacts chemically with many of the drugs and adjuvants which are incorporated into liquid preparations.

The use of hexachlorophene, a germicidal agent which is effective mainly against gram-positive organisms, is restricted to those preparations which are intended for external use only. Several years ago, an incorrectly formulated baby powder (which was found to contain 6.5% hexachlorophene) was responsible for the deaths of 30 French infants. Because of this and other evidence it can be used as a preservative only if its concentration in the final product is 0.1% or less. However, certain liquid preparations (eg, Hexachlorophene Liquid Soap USP-0.25%) are available.

Organic mercury compounds are powerful biostatic agents. Their activity may be reduced in the presence of anionic emulsifying or suspending agents. They are not suitable for oral consumption but are used at the 0.005% concentration level in ophthalmic, nasal and topical preparations.

Benzoic acid is effective only at pH 4 or less. Its solubility in certain aqueous preparations is poor and, in those instances, sodium benzoate may be used. Sorbic acid has a broad range of antimycotic activity but its antibacterial properties are more limited. It is effective only at a pH of less than 5.

Quaternary ammonium surface-active agents, eg, benzalkonium chloride, exhibit an objectionable taste and have been reported to be incompatible with a number of anionic substances. In concentrations of 1:5000 to 1:20,000 they are used in ophthalmic preparations.

3-Phenylpropan-1-ol (hydrocinnamyl alcohol) is claimed to be more effective than 2-phenylethanol and benzyl alcohol in inhibiting the growth of *P aeruginosa*, and it has been suggested that this substance may be a suitable preservative for oral suspensions and mixtures.

The methyl and propyl esters of *p*-hydroxybenzoic acid (the parabens) are used widely in the pharmaceutical industry. They are effective over a wide pH range (from about 3 to 9) and are employed up to about the 0.2% concentration level. The two esters often are used in combination in the same preparation. This achieves a higher total concentration and the mixture is active against a wide range of organisms. The hydroxybenzoates are effective against most organisms; however, their activity may be reduced in the presence of nonionic surface-active agents because of binding.

It now should be obvious that when the pharmacist dispenses or compounds the various liquid preparations he assumes responsibility, with the manufacturer, for the maintenance of product stability. The USP includes a section on stability considerations in dispensing, which should be studied in detail. Certain points are self-evident. Stock should be rotated and replaced if expiration dates on the label so indicate. Products should be stored in the manner indicated in the compendium; eg, in a cool place or a tight, light-resistant container. Further, products should be checked for evidence of instability. With respect to solutions, elixirs, and syrups, color change, precipitation and evidence of microbial or chemical gas formation are major signs of instability. Emulsions may cream but if they break (ie, there is a

separation of an oil phase) the product is considered to be unstable. Sedimentation and caking are primary indications of instability in suspensions. The presence of large particles may mean that excessive crystal growth has occurred.

The USP states that repackaging is inadvisable. However, if the product must be repackaged, care and the container specified by the compendium must be used. For example, a plastic container should never be used if a light-resistant container is specified. If a product is diluted, or where two products are mixed, the pharmacist should use his knowledge to guard against incompatibility and instability. Oral antibiotic preparations constituted into liquid form should never be mixed with other products. Since the chemical stability of extemporaneously prepared liquid preparations often is unknown, their use should be minimized and every care taken to insure that product characteristics will not change during the time it must be used by the patient.

Because of the number of excipients and additives in these preparations, it is recommended that all the ingredients be listed on the container to reduce the risks which confront hypersensitive patients when these products are administered.

Solutions

Aqueous Solutions

A solution is a homogeneous mixture that is prepared by dissolving a solid, liquid or gas in another liquid and represents a group of preparations in which the molecules of the solute or dissolved substance are dispersed among those of the solvent. Solutions also may be classified on the basis of physical or chemical properties, method of preparation, use, physical state, number of ingredients and particle size. The narrower definition herein limits the solvent to water and excludes those preparations that are sweet and/or viscid in character. This section includes, therefore, those pharmaceutical forms that are designated as *Water*, *Aromatic Waters*, *Aqueous Acids*, *Solutions*, *Douches*, *Enemas*, *Gargles*, *Mouthwashes*, *Juices*, *Nasal Solutions*, *Otic Solutions* and *Irrigation Solutions*.

Water

The major ingredient in most of the dosage forms described herein is water. It is used both as a vehicle and as a solvent for the desired flavoring or medicinal ingredients. Its tastelessness, freedom from irritating qualities and lack of pharmacological activity make it ideal for such purposes. There is, however, a tendency to assume that its purity is constant and that it can be stored, handled and used with a minimum of care. While it is true that municipal supplies must comply with Environmental Protection Agency (EPA) regulations (or comparable regulations in other countries), drinking water *must* be repurified before it can be used in pharmaceuticals. For further information on water, see Chapter 21.

Five of the six solvent waters described in the USP are used in the preparation of parenterals, irrigations or inhalations. *Purified water* must be used for all other pharmaceutical operations and, as needed, in all USP tests and assays. It must meet rigid specifications for chemical purity. Such water may be prepared by distillation, by use of ion-exchange resins or by reverse osmosis.

A wide variety of commercially available stills are used to produce distilled water. The end use of the product dictates the size of the still and extent of pretreatment of the drinking water introduced into the system. A description of stills is provided in Chapter 84. Such water may be sterile provided the condenser is sterile, but to be called sterile it must be subjected to a satisfactory sterilization process. However, it has been shown that *P aeruginosa* (and other microorganisms) can grow in the distilled water produced in hospitals. The implications of this are obvious. Sterile water may be sterile at the time of production but may lose this characteristic if it is stored improperly. Hickman *et al*,[6] by regrouping the components of conventional distillation equipment, have described a method for the continuous supply of sterile, ultrapure water. Quality-control procedures for monitoring the microbiological quality of water should be performed in the pharmaceutical manufacturer's production facilities.

The major impurities in water are calcium, iron, magnesium, manganese, silica and sodium. The cations usually are combined with the bicarbonate, sulfate or chloride anions. "Hard" waters are those that contain calcium and magnesium cations. Bicarbonates are the major impurity in "alkaline" waters.

Ion-exchange (deionization, demineralization) processes will remove most of the major impurities in water efficiently and economically. A cation exchanger, H_2R, first converts bicarbonates, sulfates and chlorides to their respective acids.

$$\left.\begin{matrix}CaSO_4\\MgSO_4\\Na_2SO_4\end{matrix}\right| + H_2R \rightarrow \left|\begin{matrix}Ca\\Mg\\Na_2\end{matrix}\right| R + H_2SO_4$$

$$\left.\begin{matrix}Ca(HCO_3)_2\\Mg(HCO_3)_2\\2NaHCO_3\end{matrix}\right| + H_2R \rightarrow \left|\begin{matrix}Ca\\Mg\\Na_2\end{matrix}\right| R + 2H_2CO_3$$

Carbonic acid decomposes to carbon dioxide (which is removed by aeration in the decarbonator) and water.

The anion exchanger may contain either a weakly basic or a strongly basic anion resin. These adsorb sulfuric, hydrochloric and nitric acids. Chemical reactions may involve complete adsorption or an exchange with some other anion.

$$H_2SO_4 + A \rightarrow A \cdot H_2SO_4$$

If the resin contains a hydroxyl radical, water is formed during the purification process.

$$H_2SO_4 + 2AOH \rightarrow A_2SO_4 + 2H_2O$$

Weakly dissociated carbonic and silicic acids can be removed only by strongly basic anion resins.

$$H_2SiO_3 + 2AOH \rightarrow A_2SiO_3 + 2H_2O$$

Unit capacity varies with the nature of the installation, but it is possible to process as much as 15,000 gal of water/min.

Deionization processes do not necessarily produce *Purified Water* which will comply with EPA requirements for drinking water. Resin columns retain phosphates and organic debris. Either alone or in combination, these substances can act as growth media for microorganisms. Observations have shown that deionized water containing 90 organisms/mL contained, after 24-hour storage, 10^6

organisms/mL. Columns can be cleaned partially of pseudomonads by recharging, but a 0.25% solution of formaldehyde will destroy most bacteria. The column must be washed thoroughly and checked for the absence of aldehyde (with a Schiffs Reagent) before it can be used to generate deionized water.

Ultraviolet radiant energy (240–280 nm), heat or filtration can be used to limit the growth, kill or remove microorganisms in water. The latter method employs membrane filters and can be used to remove bacteria from heat-labile materials as described under membrane filters in Chapter 78.

The phenomenon of osmosis involves the passage of water from a dilute solution across a semipermeable membrane to a more concentrated solution. Flow of water can be stopped by applying pressure, equal to the osmotic pressure, to the concentrated solution. The flow of water can be reversed by applying a pressure, greater than the osmotic pressure. The process of reverse osmosis utilizes the latter principle; by applying pressure, greater than the osmotic pressure, to the concentrated solution, eg, tap water, pure water may be obtained (see *Reverse Osmosis* in Chapter 77).

Cellulose acetate is used in the manufacture of semipermeable membranes for purifying water by reverse osmosis. This polymer has functional groups that can hydrogen-bond to water or other substances such as alcohol. The water molecules which enter the polymer are transported from one bonding site to the next under pressure. Because of the thin layer of pure water strongly adsorbed at the surface of the membrane, salts, to a large extent, are repelled from the surface, the higher-valent ions being repelled to a greater extent, thus causing a separation of ions from the water. Organic molecules are rejected on the basis of a sieve mechanism related to their size and shape. Small organic molecules, with a molecular weight smaller than approximately 200, will pass through the membrane material. Since there are few organic molecules with a molecular weight of less than 200 in the municipal water supply, reverse osmosis usually is sufficient for the removal of organic material. The pore sizes of the selectively permeable reverse-osmosis membranes are between 5 and 100 Å. Viruses and bacteria larger than 100 Å are rejected if no imperfections exist in the membrane. The membranes may and do develop openings which permit the passage of microorganisms. Because of the semistatic conditions, bacteria can grow both upstream and downstream of the membrane. Improvements in membranes are being made continually in type and manufacturing process such as the use of polyamide materials. It is expected that the preparation of water with negligible or no bacteria present will be achieved by this process.

The selection of water-treatment equipment depends upon the quality of water to be tested, the quality of water required and the specific pharmaceutical purpose of the water. Frequently, two or more methods are used to produce the water desired, for example, filtration and distillation, or filtration, reverse osmosis and ion exchange.

Aromatic Waters

Aromatic waters, known also as medicated waters, are clear, saturated aqueous solutions of volatile oils or other aromatic or volatile substances. Their odors and tastes are similar to those of the drugs or volatile substances from which they are prepared, and the preparations should be free from empyreumatic (smoke-like) and other foreign odors. They are used principally as flavored or perfumed vehicles. The volatile substances from which they are to be made should be of pharmacopeial quality or, in the case of nonofficial preparations, of the best quality if the finest flavors are to be obtained.

Aromatic waters may be prepared by one of two official processes.

Distillation—Different authorities give different directions for preparing aromatic waters by distillation. For fresh drugs the proportions range from 1 part of drug to 2 of distillate, to 2 parts of drug to 1 of distillate. For dried drugs such as cinnamon, anise, dill, caraway and fennel the proportion is 1 part of drug to 10 of distillate. For dried leaf drugs such as peppermint the proportion is 3 parts of drug to 10 of distillate. The drug should be contused or coarsely ground and combined with a sufficient quantity of *Purified Water*. Most of the water then is distilled; care should be taken to avoid charring or scorching the substances to prevent the formation of empyreumatic odors. On completion of the distillation, any excess oil in the distillate is removed and, if necessary, the clear-water portion is filtered.

Solution—Aromatic waters may be prepared by shaking repeatedly 2 g or (2 mL if a liquid) of the volatile substance with 1000 mL of purified water for 15 min. The mixture is set aside for 12 hr, filtered through wetted filter paper and made to volume (1000 mL) by adding purified water through the filter. Peppermint Water USP can be prepared by either of the two official methods.

Alternately aromatic waters also may be prepared by incorporating thoroughly the volatile oil with 15 g of talc, or with a sufficient quantity of purified siliceous earth or pulped filter paper. Purified water (1000 mL) is added and the mixture is agitated for 10 min. The water then is filtered (and, if necessary, refiltered) and its volume adjusted to 1000 mL by passing purified water through the filter.

This is the process most frequently employed since the water can be prepared promptly, only 10 minutes of agitation being required. The use of talc, purified siliceous earth or pulped filter paper greatly increases the surface of the volatile substance, insuring more rapid saturation of the water. These dispersing substances also form an efficient filter bed which produces a clear solution. They also are unreactive.

Other methods have been suggested for preparing aromatic waters based on the use of soluble concentrates or on incorporation of solubilizing agents such as polysorbate 20 (Tween 20, *Atlas*). However, such preparations are susceptible to mold growth and, in concentrations higher than 2%, impart an objectionable oily taste.

Concentrated waters (eg, peppermint, dill, cinnamon, caraway and anise) may be prepared as follows:

Dissolve 20 mL of the volatile oil in 600 mL of 90% ethanol. Add sufficient purified water in successive small portions to produce 1000 mL. Shake vigorously after each addition. Add 50 g of sterilized purified talc, shake occasionally for several hours and filter.

If anise concentrate is being prepared, the volume of ethanol must be increased to 700 mL.

The aromatic water is prepared by diluting the concentrate with 39 times its volume of water. In general, these methods yield aromatic waters that are slightly inferior in quality to those prepared by distillation or solution.

The chemical composition of many of the volatile oils used in preparing pharmaceuticals and cosmetics now is known. Similarly, many synthetic aromatic substances have a characteristic odor; eg, geranyl phenyl acetate has a honey odor. Such substances, either alone or in combination, can be used in nonofficial preparations and, by combining them in definite proportions, it is possible to produce substitutes for the officially recognized oil. Imitation Otto of Rose (which contains phenylethyl alcohol, rhodinol, citronellol and other ingredients) is an example of the types of substitutes which are now available. Additional information regarding the appropriate preparation of aromatic waters is provided in RPS-17, Chapter 84.

Incompatibilities—The principal difficulty experienced in compounding prescriptions containing aromatic waters is due to a "salting out" action of certain ingredients, such as very soluble salts, on the volatile principle of the aromatic water. A replacement of part of the aromatic water with purified water is permissible when no other function is being

served than that of a vehicle. Otherwise, a dilution of the product, with a suitable increase in dosage, is indicated.

Preservation—Aromatic waters will deteriorate with time and should, therefore, be made in small quantities and protected from intense light, excessive heat and stored in airtight, light-resistant containers. Deterioration may be due to volatilization, decomposition or mold growth and will produce solutions that are cloudy and have lost all traces of their agreeable odor. Distilled water usually is contaminated with mold-producing organisms. *Recently* distilled and boiled water should, therefore, be used in the preparation of medicated waters. No preservative should be added to medicated waters. If they become cloudy or otherwise deteriorate, they should be discarded.

Aqueous Acids

The official inorganic acids and certain organic acids, although of minor significance as therapeutic agents, are of great importance in chemical and pharmaceutical manufacturing. This is especially true of acetic, hydrochloric and nitric acids.

Percentage Strengths—Many of the more important inorganic acids are available commercially in the form of concentrated aqueous solutions. The percentage strength varies from one acid to another and depends on the solubility and stability of the solute in water and on the manufacturing process. Thus, the official Hydrochloric Acid contains from 36.5 to 38% by weight of HCl, whereas Nitric Acid contains from 69 to 71% by weight of HNO_3.

Because the strengths of these concentrated acids are stated in terms of % by weight, it is essential that specific gravities also be provided if one is to be able to calculate conveniently the amount of absolute acid contained in a unit volume of the solution as purchased. The mathematical relationship involved is given by the equation $M = V \times S \times F$, where M is the mass in g of absolute acid contained in V mL of solution having a specific gravity S and a fractional percentage strength F. As an example, Hydrochloric Acid containing 36.93% by weight of HCl has a specific gravity of 1.1875. Therefore, the amount of absolute HCl supplied by 100 mL of this solution is given by:

$$M = 100 \times 1.1875 \times 0.3693 = 43.85 \text{ g HCl}$$

Incompatibilities—Although many of the reactions characteristic of acids offer opportunities for incompatibilities, only a few are of sufficient importance to require more than casual mention. Acids and acid salts decompose carbonates with liberation of carbon dioxide and, in a closed container, sufficient pressure may be developed to produce an explosion. Inorganic acids react with salts of organic acids to produce the free organic acid and a salt of the inorganic acid. If insoluble, the organic acid will be precipitated. Thus, salicylic acid and benzoic acid are precipitated from solutions of salicylates and benzoates. Boric acid likewise is precipitated from concentrated solutions of borates. By a similar reaction, certain soluble organic compounds are converted into an insoluble form. Phenobarbital sodium, for example, is converted into phenobarbital which will precipitate in aqueous solution.

The ability of acids to combine with alkaloids and other organic compounds containing a basic nitrogen atom is used in preparing soluble salts of these substances.

It should be borne in mind that certain solutions, syrups, elixirs and other pharmaceutical preparations, may contain free acid, which causes these preparations to exhibit the incompatibilities characteristic of the acid.

Acids also possess the incompatibilities of the anions which they contain and, in the case of organic acids, these are frequently of prime importance. These are discussed under the specific anions.

Diluted Acids—The diluted acids in the USP are aqueous solutions of acids, of a suitable strength (usually 10% *w/v* but Diluted Acetic Acid is 6% *w/v*) for internal administration or for the manufacture of other preparations.

The strengths of the official undiluted acids are expressed as percentages *w/w*, whereas the strengths of the official diluted acids are expressed as percent *w/v*. It, therefore, becomes necessary to consider the specific gravities of the concentrated acids when calculating the volume required to make a given quantity of diluted acid. The following equation will give the number of mL required to make 1000 mL of diluted acid:

$$\frac{\text{Strength of diluted acid} \times 1000}{\text{Strength of undiluted acid} \times \text{sp gr of undiluted acid}}$$

Thus, if one wishes to make 1000 mL of Diluted Hydrochloric Acid USP using Hydrochloric Acid which assays 37.5% HCl (sp *gr* 1.18), the amount required is

$$\frac{10 \times 1000}{37.5 \times 1.18} = 226 \text{ mL}$$

Diluted Hydrochloric Acid USP is used in the treatment of achlorhydria. However, it may irritate the mucous membrane of the mouth and attack the enamel of the teeth. The usual dose is 5 mL, well-diluted with water. In the treatment of achlorhydria no attempt is made to administer more than a relief-producing dose. The normal pH of the gastric juice is 0.9 to 1.5 and, in order to attain this level, particularly in severe cases of gastric malfunction, somewhat larger doses of the acid would be required.

Solutions

A solution is a liquid preparation that contains one or more soluble chemical substances dissolved in water. The solute usually is nonvolatile. Solutions are used for the specific therapeutic effect of the solute, either internally or externally. Although the emphasis here is on the aqueous solution, certain preparations of this type (syrups, infusions and decoctions) have distinctive characteristics and, therefore, are described later in the chapter.

Solvents, solubility and general methods for the incorporation of a solute in a solvent are discussed in Chapter 16. Solutions are usually bottled automatically with equipment of the type shown in Fig 83-1.

Preparation—A specific method of preparation is given in the compendia for most solutions. These procedures fall into three main categories.

Simple Solutions—Solutions of this type are prepared by dissolving the solute in a suitable solvent. The solvent may contain other ingredients which stabilize or solubilize the active ingredient. Calcium Hydroxide Topical Solution (Lime Water), Sodium Phosphates Oral Solution and Strong Iodine Solution are examples.

Calcium Hydroxide Topical Solution contains, in each 100 mL, not less than 140 mg of $Ca(OH)_2$. The solution is prepared by agitating vigorously 3 g of calcium hydroxide with 1000 mL of cool, purified water. Excess calcium hydroxide is allowed to settle out and the clear, supernatant liquid dispensed.

An increase in solvent temperature usually implies an increase in solute solubility. This rule does not apply, however, to the solubility of calcium hydroxide in water, which decreases with increasing temperature. The official solution is prepared at 25°.

Solutions containing hydroxides react with the carbon dioxide in the atmosphere.

$$OH^- + CO_2 \rightarrow HCO_3^-$$

$$OH^- + HCO_3^- \rightarrow CO_3^{2-} + H_2O$$

$$Ca^{2+} + CO_3^{2-} \rightarrow CaCO_3$$

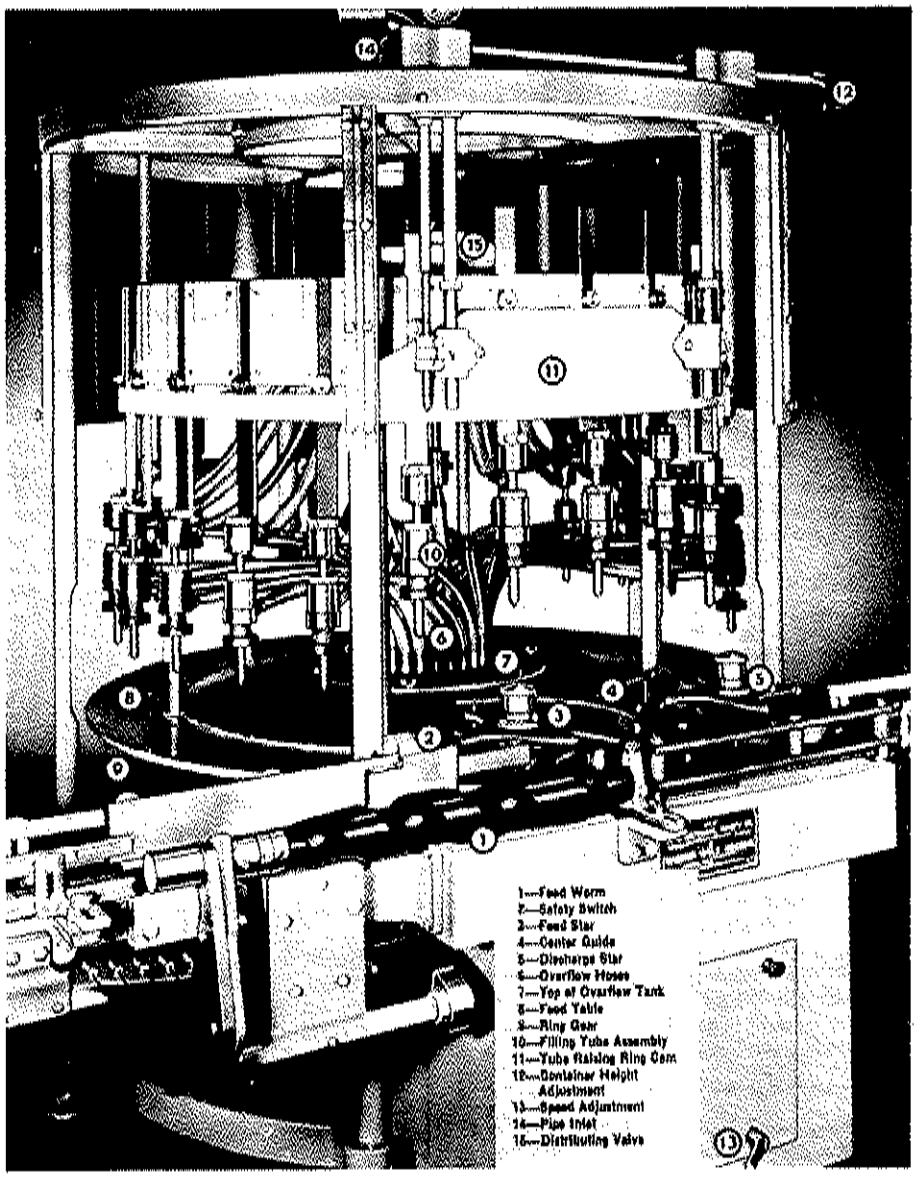


Fig 83-1. A rotary gravity bottle filler (courtesy, US Bottlers).

Calcium Hydroxide Topical Solution, therefore, should be preserved in well-filled, tight containers, at a temperature not exceeding 25°.

Strong Iodine Solution contains, in each 100 mL, 4.5–5.5 g of iodine, and 9.5–10.5 g of potassium iodide. It is prepared by dissolving 50 g of iodine in 100 mL of purified water containing 100 g of potassium iodide. Sufficient purified water then is added to make 1000 mL of solution.

One g of iodine dissolves in 2950 mL of water. However, solutions of iodides dissolve large quantities of iodine. Strong Iodine Solution is, therefore, a solution of polyiodides in excess iodide.

$$I^- + nI_2 \rightarrow I^-_{(2n+1)}$$

Doubly charged anions may be found also

$$2I^- + nI_2 \rightarrow I^{2-}_{(2n+2)}$$

Strong Iodine Solution is classified as an antigoitrogenic. The usual dose is 0.3 mL, 3 times a day.

Several antibiotics (eg, cloxacillin sodium, nafcillin sodium and vancomycin), because they are relatively unstable in aqueous solution, are prepared by manufacturers as dry powders or granules in combination with suitable buffers, colors, diluents, dispersants, flavors and/or preservatives. These preparations, Cloxacillin Sodium for Oral Solution, Nafcillin for Oral Solution and Vancomycin for Oral Solution meet the requirements of the USP. Upon dispensing to the patient, the pharmacist adds the appropriate amount of water. The products are stable for up to 14 days when refrigerated. This period usually provides sufficient time for the patient to complete the administration of all the medication.

Solution by Chemical Reaction—These solutions are prepared by reacting two or more solutes with each other in a suitable solvent. An example is Aluminum Subacetate Topical Solution.

Aluminum sulfate (145 g) is dissolved in 600 mL of cold water. The solution is filtered, and precipitated calcium carbonate (70 g) is added, in several portions, with constant stirring. Acetic acid (160 mL) is added slowly and the mixture set aside for 24 hr. The product is filtered and the magma on the Buchner filter washed with cold water until the total filtrate measures 1000 mL.

The solution contains pentaquohydroxo- and tetraquodihydroxoaluminum (III) acetates and sulfates dissolved in an aqueous medium saturated with calcium sulfate. The solution contains a small amount of acetic acid. It is stabilized by the addition of not more than 0.9% boric acid.

The reactions involved in the preparation of the solution are given below. The hexaquo aluminum cations first are converted to the nonirritating $[Al(H_2O)_5(OH)]^{2+}$ and $[Al(H_2O)_4(OH)_2]^{1+}$ cations.

$$[Al(H_2O)_6]^{3+} + CO_3^{2-} \rightarrow [Al(H_2O)_5(OH)]^{2+} + HCO_3^-$$

$$[Al(H_2O)_6]^{3+} + HCO_3^- \rightarrow [Al(H_2O)_5(OH)]^{2+} + H_2O + CO_2$$

As the concentration of the hexaquo cations decreases, secondary reactions involving carbonate and bicarbonate occur.

$$[Al(H_2O)_5(OH)]^{2+} CO_3^{2-} \rightarrow [Al(H_2O)_4(OH)_2]^+ + HCO_3^-$$

$$[Al(H_2O)_5(OH)]^{2+} + HCO_3^- \rightarrow [Al(H_2O)_4(OH)_2]^+ + H_2CO_3$$

The pH of the solution now favors the precipitation of dissolved calcium ions as the insoluble sulfate. Acetic acid now is added. The bicarbonate which is formed in the final stages of the procedure is removed as carbon dioxide.

Aluminum Subacetate Topical Solution is used in the preparation of Aluminum Acetate Topical Solution (Burow's Solution). The latter solution contains 15 mL of glacial acetic acid, 845 mL of Aluminum Subacetate Topical Solution and sufficient water to make 1000 mL. It is defined as a solution of aluminum acetate in approximately 5%, by weight, of acetic acid in water. It is stabilized by the addition of not more than 0.5% boric acid.

Solution by Extraction—Drugs or pharmaceutical necessities of vegetable or animal origin often are extracted with water or with water containing other substances. Preparations of this type may be classified as solutions but, more often, are classified as extracts.

Douches

A douche is an aqueous solution directed against a part or into a cavity of the body. It functions as a cleansing or antiseptic agent. An *eye douche*, used to remove foreign particles and discharges from the eyes, is directed gently at an oblique angle and allowed to run from the inner to the outer corner of the eye. *Pharyngeal douches* are used to prepare the interior of the throat for an operation and cleanse it in suppurative conditions. Similarly, there are *nasal douches* and *vaginal douches*. Douches usually are directed to the appropriate body part by using bulb syringes (Chapter 104).

Douches most frequently are dispensed in the form of a powder with directions for dissolving in a specified quantity of water (usually warm). However, tablets for preparing solutions are available (eg, Dobell's Solution Tablets) or the solution may be prepared by the pharmacist. If powders or tablets are supplied, they must be free from insoluble material, in order to produce a clear solution. Tablets are produced by the usual processes (see Chapter 89) but any lubricants or diluents used must be readily soluble in water. Boric acid may be used as a lubricant and sodium chloride normally is used as a diluent. Tablets deteriorate on exposure to moist air and should be stored in airtight containers.

Preparations of this type may contain alum, zinc sulfate, boric acid, phenol or sodium borate. The ingredients in one douche are alum (4 g), zinc sulfate (4 g), liquefied phenol (5 mL), glycerin (125 mL) and water (qs to make 1000 mL of solution). Sodium borate (borax, sodium tetraborate) is used in the preparation of Compound Sodium Borate Solution NF XI (Dobell's Solution). Its aqueous solution is alkaline to litmus paper. In the presence of water, sodium metaborate, boric acid and sodium hydroxide are formed.

$$Na_2B_4O_7 + 3H_2O \rightarrow 2NaBO_2 + 2H_3BO_3$$

$$NaBO_2 + 2H_2O \rightarrow NaOH + H_3BO_3$$

The official solution contains sodium borate, sodium bicarbonate, liquefied phenol and glycerin. The reaction between boric acid and glycerin is given in the section on *Mouthwashes*. See also the section on *Honeys* for a discussion on the toxic manifestations associated with the topical application of boric acid and borax.

Douches are not official as a class of preparations but several substances in the compendia frequently are employed as such in weak solutions, eg, benzalkonium chloride is used in various douches and Compound Sodium Borate Solution is used as a nasal or pharyngeal douche. A sodium bicarbonate vaginal douche has been used to improve the postcoital test.

Vaginal douches are used for cleansing the vagina and hygienic purposes. Liquid concentrates or powders, which may be prepared in bulk or as single-use packages, should be diluted or dissolved in the appropriate amount of warm water prior to use. The ingredients used in vaginal douches include antimicrobial agents such as benzalkonium chloride, the parabens or chlorothymol, anesthetics or antipruritics such as phenol or menthol, Astringents such as zinc sulfate or potassium alum, surface-active agents such as sodium lauryl sulfate and chemicals to alter the pH such as sodium bicarbonate or citric acid also are used.

Enemas

These preparations are rectal injections employed to evacuate the bowel (evacuation enemas), influence the general system by absorption (retention enemas) or to affect locally the seat of disease. They may possess anthelmintic, nutritive, sedative or stimulating properties, or they may contain radiopaque substances for roentgenographic examination of the lower bowel. Some official retention enemas are those of aminophylline, hydrocortisone and methylprednisolone acetate. Since they are to be retained in the intestine, they should not be used in larger quantities than 150 mL for an adult. Usually, the volume is considerably smaller, such as a few mL. *Microenema* is a term used to describe these small-volume preparations. Vehicles for retention microenemas have been formulated with small quantities of ethanol and propylene glycol, and no significant difference in irritation, as compared with water, was found. A number of drugs such as valproic acid, indomethacin and metronidazole have been formulated as microenemas for the purpose of absorption. The absorption of large molecular weight drugs, such as insulin, is under current investigation.

Starch enema may be used either by itself or as a vehicle for other forms of medication. A thin paste is made by triturating 30 g of powdered starch with 200 mL of cold water. Sufficient boiling water is added to make 1000 mL of enema. The preparation then is reheated to obtain a transparent liquid.

Sodium chloride, sodium bicarbonate, sodium monohydrogen phosphate and sodium dihydrogen phosphate are used in enemas to evacuate the bowel. These substances may be used alone, in combination with each other or in combination with irritants such as soap. Enema of Soap BPC 1963 is prepared by dissolving 50 g of soft soap in sufficient purified water to make 1000 mL of enema. Fleet Enema, a commercially available enema containing 16 g of sodium acid phosphate and 6 g of sodium phosphate in 100 mL, is marketed as a single-dose disposable unit. Evacuation enemas usually are given at body temperature in quantities of 1 to 2 pt injected slowly with a syringe.

Sulfasalazine rectal enema has been administered for the treatment of ulcerative colitis and may be prepared by dispersing the tablets (1-g strength) in 250 mL water. Barium sulfate enema contains 120 g of barium sulfate, 100 mL of acacia mucilage and sufficient starch enema to make 500 mL.

Gargles

Gargles are aqueous solutions used for treating the pharynx and nasopharynx by forcing air from the lungs through

the gargle which is held in the throat. Many gargles must be diluted with water prior to use. Although mouthwashes are considered as a separate class of pharmaceuticals, many are used as gargles, either as is, or diluted with water.

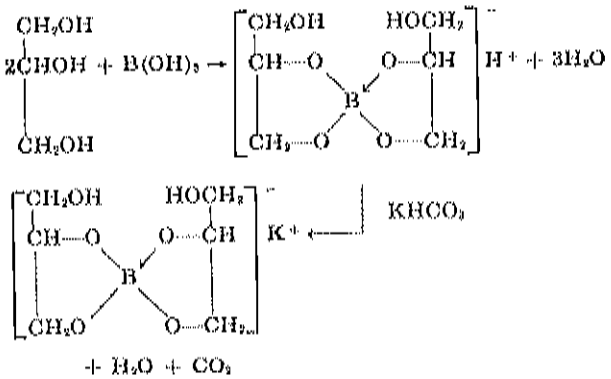
Potassium Chlorate and Phenol Gargle is official in the PC. It contains potassium chlorate, 30 g, patent blue V (Color Index No 42051) commercial food grade (0.01 g), liquified phenol (15 mL) and water for preparations qs to 1000 mL. It should be diluted with 10 volumes of warm water before use. The product should be labeled so that it cannot be mistaken for preparations intended for internal administration.

A flavored solution containing 7.5% povidone-iodine and 35% alcohol (*Isodine*) is available commercially as a mouthwash or gargle after suitable dilution.

Mouthwashes

A mouthwash is an aqueous solution which is most often used for its deodorant, refreshing or antiseptic effect or for control of plaque. It may contain alcohol, glycerin, synthetic sweeteners and surface-active, flavoring and coloring agents. Commercial preparations contain such local anti-infective agents as hexetidine and cetylpyridinium chloride. They may be either acidic or basic in reaction and, in some instances, may be effective in reducing bacterial concentrations and odors in the mouth for short periods of time.

The products of commerce (eg, Cepacol, Listerine, Micrin or Scope) vary widely in composition. Compound Sodium Borate Solution NF XI (Dobell's Solution) is used as an antiseptic mouthwash and gargle. Antiseptic Solution and Mouthwash are described in NF XII. The latter wash contains sodium borate, glycerin and potassium bicarbonate. The reactions which take place when these substances are dissolved in water are given below.



Compound Sodium Chloride Mouthwash, and Zinc Sulphate and Zinc Chloride Mouthwash are described in the BPC. The former wash contains sodium chloride, sodium bicarbonate, concentrated peppermint emulsion and double-strength chloroform water.

Mouthwashes may be used for a number of purposes: for example, cetylpyridinum chloride and dibucaine hydrochloric mouthwashes provide satisfactory relief of pain in patients with ulcerative lesions of the mouth, mouthwashes or creams containing carbenoxolone are highly effective dosage forms for the treatment of orofacial herpes simplex infections, and undetected oral cancer has been recognized using toluidine blue in the form of a mouth rinse.

Juices

A juice is prepared from fresh ripe fruit, is aqueous in character and is used in making syrups which are employed as vehicles. The freshly expressed juice is preserved with benzoic acid and allowed to stand at room temperature for several days, until the pectins which naturally are present are destroyed by enzymatic action, as indicated by the filtered juice yielding a clear solution with alcohol. Pectins, if allowed to remain, would cause precipitation in the final syrup.

Cherry Juice is described in the current USP and Raspberry Juice in USP XVIII. Concentrated Raspberry Juice PC is prepared from the clarified juice of raspberries. Pectinase is stirred into pulped raspberries and the mixture allowed to stand for 12 hours. The pulp is pressed, the juice clarified and sufficient sucrose added to adjust the weight at 20° to 1.050 to 1.060 g per mL. The juice then is concentrated to one-sixth of its original volume. Sufficient sulfurous acid or sodium metabisulfite is added as a preservative.

Artificial flavors now have replaced many of the natural fruit juices. Although they lack the flavor of the natural juice, they are more stable and easier to incorporate into the final pharmaceutical form.

Recent information on cranberry juice indicates that it may be effective in controlling some urinary tract infections and urolithiosis.

Nasal Solutions

Nasal solutions are usually aqueous solutions designed to be administered to the nasal passages in drops or sprays. While many of the drugs are administered for their local sympathomimetic effect such as Ephedrine Sulfate or Naphazoline Hydrochloride Nasal Solution USP, to reduce nasal congestion, a few other official preparations, Lypressin Nasal Solution USP and Oxytocin Nasal Solution USP, are administered in spray form for their systemic effect for the treatment of diabetes insipidus and *milk letdown* prior to breast feeding, respectively. The current route of administration of peptides and proteins is limited to parenteral injection because of inactivation within the gastrointestinal tract. As a result, there is considerable research on intranasal delivery of these drugs such as analogs of enkephalins or lutenizing hormone releasing hormone and insulin.

Nasal solutions are prepared so that they are similar in many respects to nasal secretions, so that normal ciliary action is maintained. Thus, the aqueous nasal solutions usually are isotonic and slightly buffered to maintain a pH of 5.5 to 6.5. In addition, antimicrobial preservatives, similar to those used in ophthalmic preparations, and appropriate drug stabilizers, if required, are included in the formulation.

Commercial nasal preparations, in addition to the drugs listed above also include antibiotics, antihistamines and drugs for asthma prophylaxis.

A formula for Ephedrine Nasal Drops PC is

Ephedrine hydrochloride	0.5 g
Chlorobutanol	0.5 g
Sodium Chloride	0.8 g
Water for preparations	to 100 mL

Current studies indicate that nasal sprays are deposited mainly in the atrium and cleared slowly into the pharynx with the patient in an upright position. Drops spread more extensively than the spray and three drops cover most of the walls of the nasal cavity, with the patient in a supine position and head tilted back and turned left and right. It is suggested that drop delivery, with appropriate movement by the patient, leads to extensive coverage of the walls of the nasal cavity.

Otic Solutions

These solutions occasionally are referred to as aural preparations. Other otic preparations often include formula-

tions such as suspensions and ointments for topical application in the ear.

The main classes of drugs used for topical administration to the ear include analgesics, eg, benzocaine; antibiotics, eg, neomycin; and anti-inflammatory agents, eg, cortisone. The USP preparations include Antipyrine and Benzocaine Otic Solution. The Neomycin and Polymyxin B Sulfates and Hydrocortisone Otic Solutions contain appropriate buffers and dispersants usually in an aqueous solution. These preparations include the main types of solvents used, namely glycerin or water. The viscous glycerin vehicle permits the drug to remain in the ear for a long time. Anhydrous glycerin, being hygroscopic, tends to remove moisture from surrounding tissues, thus reducing swelling. Viscous liquids like glycerin or propylene glycol either are used alone or in combination with a surfactant to aid in the removal of cerumen (ear wax).

In order to provide sufficient time for aqueous preparations to act, it is necessary for the patient to remain on his side for a few minutes so the drops do not run out of the ear. Otic preparations are dispensed in a container which permits the administration of drops.

Irrigation Solutions

These solutions are used to wash or bathe surgical incisions, wounds or body tissues. Because they come in contact with exposed tissue, they must meet stringent requirements for Injections of the USP such as sterility, particulate matter and the requirements of the Pyrogen Test. These products are prepared by dissolving the active ingredient in Water for Injection. They are packaged in single-dose containers, preferably Type I or Type II glass, or suitable plastic containers, and then sterilized. See Chapter 78 for sterilization procedures. A number of irrigations are described in the USP: Acetic Acid Irrigation for bladder irrigation, Aminoacetic Acid Irrigation for urethral surgery and Sodium Chloride Irrigation for washing wounds. Other drugs such as amphotericin B also may be formulated as irrigations.

Sweet or Other Viscid Aqueous Solutions

Solutions which are sweet or viscid include syrups, honeys, mucilages and jellies. All of these are viscous liquids or semisolids. The basic sweet or viscid substances giving body to these preparations are sugars, polyols or polysaccharides (gums).

Syrups

Syrups are concentrated solutions of sugar such as sucrose in water or other aqueous liquid. When Purified Water alone is used in making the solution of sucrose, the preparation is known as *Syrup*, or *simple syrup*. In addition to sucrose, certain other polyols, such as glycerin or sorbitol, may be added to retard crystallization of sucrose or to increase the solubility of added ingredients. Alcohol often is included as a preservative and also as a solvent for flavors; further resistance to microbial attack can be enhanced by incorporating antimicrobial agents. When the aqueous preparation contains some added medicinal substance, the syrup is called a *medicated syrup*. A *flavored* syrup is one which usually is not medicated, but which contains various aromatic or pleasantly flavored substances and is intended to be used as a vehicle or flavor for prescriptions.

Flavored syrups offer unusual opportunities as vehicles in extemporaneous compounding and are accepted readily by both children and adults. Because they contain no, or very little, alcohol they are vehicles of choice for many of the drugs that are prescribed by pediatricians. Their lack of alcohol makes them superior solvents for water-soluble substances. However, sucrose-based medicines continuously administered to children apparently cause an increase in dental caries and gingivitis; consequently, alternate formulations of the drug either unsweetened or sweetened with noncariogenic substances should be considered. A knowledge of the sugar content of liquid medicines is useful for patients who are on a restricted caloric intake; a list has been prepared by Bergen.[6]

Syrups possess remarkable masking properties for bitter or saline drugs. Glycyrrhiza syrup has been recommended for disguising the salty taste of bromides, iodides and chlorides. This has been attributed to its colloidal character and its double sweetness—the immediate sweetness of the sugar and the lingering sweetness of the glycyrrhizin. This syrup is also of value in masking bitterness in preparations containing the B complex vitamins. Acacia Syrup USP, because of its colloidal character, is of particular value as a vehicle for masking the disagreeable taste of many medicaments. Raspberry Syrup PC is one of the most efficient flavoring agents and is especially useful in masking the taste of bitter drugs. Many factors, however, enter into the choice of a suitable flavoring agent. Literature reports are often contradictory and there appears to be no substitute for the taste panel. The literature on this subject has been reviewed by Meer,[7] and this reference and Chapter 66 should be consulted for further information on the flavoring of pharmaceuticals and the preparation of a number of official syrups. A series of papers by Schumacher deals with improving the palatability of bulk-compounded products using flavoring and sweetening agents.[8]

In manufacturing syrups the sucrose must be selected carefully and a purified water, free from foreign substances, and clean vessels and containers must be used. The operation must be conducted with care to avoid contamination, if the products are to be stable.

It is important that the concentration of sucrose approach but not quite reach the saturation point. In dilute solutions sucrose provides an excellent nutrient for molds, yeasts and other microorganisms. In concentrations of 65% by weight or more, the solution will retard the growth of such microorganisms. However, a saturated solution may lead to crystallization of a part of the sucrose under conditions of changing temperature.

When heat is used in the preparation of syrups, there is almost certain to be an inversion of a slight portion of the sucrose. Sucrose solutions are dextrorotary but, as hydrolysis proceeds, the optical rotation decreases and becomes negative when the reaction is complete. This reaction is termed *inversion* because *invert sugar* (dextrose plus levulose) is formed. The speed of inversion is increased greatly by the presence of acids; the hydrogen ion acts as a catalyst in this hydrolytic reaction. Invert sugar is more readily fermentable than sucrose and tends to be darker in color. Nevertheless, its two reducing sugars are of value in retarding the oxidation of other substances.

Invert Syrup is described in the PC. It is prepared by hydrolyzing sucrose with hydrochloric acid and neutralizing the solution with calcium or sodium carbonate. The sucrose in the 66.7% *w/w* solution must be at least 95% inverted. The monograph states that invert syrup, when mixed in suitable proportions with syrup, prevents the deposition of crystals of sucrose under most conditions of storage.

The levulose formed during inversion is sweeter than su-

crose and, therefore, the resulting syrup is sweeter than the original syrup. The relative sweetness of levulose, sucrose and dextrose is in the ratio of 173:100:74. Thus, invert sugar is 1/100 (173 + 74)½ = 1.23 times as sweet as sucrose. The levulose formed during the hydrolysis also is responsible for the darkening of syrup. It is sensitive to heat and darkens readily, particularly in solution. When syrup or sucrose is overheated, it caramelizes. See *Caramel* (page 1290). Occasionally, it is appropriate to use a sugar-free liquid preparation; a list of these has been published.[9]

Preparation—Syrups are prepared in various ways, the choice of the proper method depending on the physical and chemical characteristics of the substances entering into the preparation.

Solution with Heat—This is the usual method of making syrups when the valuable constituent is neither volatile nor injured by heat, and when it is desirable to make the syrup rapidly. The sucrose usually is added to the purified water or aqueous solution and heated until solution is effected, then strained and sufficient purified water added to make the desired weight or volume. If the syrup is made from an infusion, a decoction or an aqueous solution containing organic matter, such as sap from maple trees, it usually is proper to heat the syrup to the boiling point to coagulate albuminous matter; subsequently, this is separated by straining. If the albumin or other impurities were permitted to remain in the syrup, fermentation probably would be induced in warm weather. Saccharometers are very useful in making syrups by the hot process in cases where the proper specific gravity of the finished syrup is known. They may be floated in the syrup while boiling, and thus the exact degree of concentration determined without waiting to cool the syrup and having to heat it again to concentrate it further. When taking a reading of the specific gravity of the hot syrup, allowance must be made for the variation from the official temperature (specific gravities in the USP are taken at 25°).

Excessive heating of syrups at the boiling temperature is undesirable since more or less inversion of the sucrose occurs with an increased tendency to ferment. Syrups cannot be sterilized in an autoclave without some caramelization. This is indicated by a yellowish or brownish color resulting from the formation of caramel by the action of heat upon sucrose.

The formula and procedure given for Acacia Syrup (page 1301) illustrates this method of preparation.

Agitation without Heat—This process is used in those cases where heat would cause the loss of valuable, volatile constituents. In making quantities up to 2000 mL, the sucrose should be added to the aqueous solution in a bottle of about twice the size required for the syrup. This permits active agitation and rapid solution. Stoppering the bottle is important, as it prevents contamination and loss during the process. The bottle should be allowed to lie on its side when not being agitated. Glass-lined tanks with mechanical agitators, especially adapted to dissolving of sucrose, are used for making syrups in large quantities.

This method and that previously described are used for the preparation of a wide variety of preparations that are popularly described as syrups. Most cough syrups, for example, contain sucrose and one or more active ingredients. However, the exact composition of such products is not given on the label. Furthermore, some of these products are listed in the USP but no directions are given for their preparation. For example, Guaifenesin Syrup USP (glyceryl guaiacolate syrup) is official but the only known ingredients are guaifenesin (glyceryl guaiacolate) and ethanol (not less than 3% or more than 4%).

The PC, on the other hand, gives a method for the preparation of Codeine Phosphate Syrup. This contains codeine phosphate (5 g), water for preparations (15 mL), chloroform spirit (25 mL) and sufficient syrup to make 1000 mL. It can be used for the relief of cough. Another syrup for this purpose is Codeine Linctus PC. This is really a medicated syrup which possesses demulcent, expectorant or sedative properties. Unlike the syrup, it is colored and flavored. The formula for Codeine Linctus PC is:

Codeine Phosphate	3 g
Compound Tartrazine Solution	10 mL
Benzoic Acid Solution	20 mL
Chloroform Spirit	20 mL
Water for Preparations	20 mL
Lemon Syrup	200 mL
Syrup	to 1000 mL

Dissolve the codeine phosphate in the water, add 500 mL of the syrup and mix. Add the other ingredients and sufficient syrup to produce 1000 mL.

For pediatric use, 200 mL of this linctus is diluted with sufficient syrup to make 1000 mL. If sugar is contraindicated in the diet, Diabetic Codeine Linctus can be used:

Codeine Phosphate	3 g
Citric Acid monohydrate	5 g
Lemon Spirit	1 mL
Compound Tartrazine Solution	10 mL
Benzoic Acid Solution	20 mL
Chloroform Spirit	20 mL
Water for Preparations	20 mL
Sorbitol Solution	to 1000 mL

Dissolve the codeine phosphate and the citric acid in the water, add 750 mL of the sorbitol solution and mix. Add the other ingredients and sufficient sorbitol solution to produce 1000 mL.

Sorbitol Solution is the sweetening agent and contains 70% *w/w* of total solids, consisting mainly of D-sorbitol. It has about half the sweetening power of syrup. In the US the FDA has banned the use of chloroform in medicines and cosmetics because of reported carcinogenicity in animals.

Basic formulations can be varied easily to produce the highly advertised articles of commerce. The prescription-only drug (eg, codeine phosphate or methadone) must, of course, be omitted from the formulation but, in certain countries, such as Canada, a decreased quantity of codeine phosphate is permitted in an OTC cough syrup. In addition to the ingredients cited or listed in the official compendia (eg, tolu, squill or ipecacuanha), many cough syrups contain an antihistamine.

Many other active ingredients (eg, ephedrine sulfate, dicyclomine hydrochloride, chloral hydrate or chlorpromazine hydrochloride) are marketed as syrups. Like cough syrups, these preparations are flavored, colored and recommended in those instances where the patient cannot swallow the solid dosage form.

Addition of a Medicating Liquid to Syrup—This method is resorted to in those cases in which fluidextracts, tinctures or other liquids are added to syrup to medicate it. Syrups made in this way usually develop precipitates since alcohol is often an ingredient of the liquids thus used, and the resinous and oily substances dissolved by the alcohol precipitate when mixed with the syrup, producing unsightly preparations. A modification of this process, frequently adopted, consists of mixing the fluidextract or tincture with the water, allowing the mixture to stand to permit the separation of insoluble constituents, filtering and then dissolving the sucrose in the filtrate. It is obvious that this procedure is not permissible when the precipitated ingredients are the valuable medicinal agents.

The formula and procedure given for Aromatic Eriodictyon Syrup USP (page 1301) illustrate this method of preparation.

Percolation—In this procedure, purified water, or an aqueous solution, is permitted to pass slowly through a bed of crystalline sucrose, thus dissolving it and forming a syrup. A cotton pledget is placed in the neck of the percolator and the water or aqueous solution added. By means of a suitable stopcock the flow is regulated so that drops appear in rapid succession. If necessary, a portion of the liquid is recycled through the percolator to dissolve all the sucrose. Finally, sufficient purified water is passed through the cotton to make the required volume.

To be successful in using this process, care in several particulars must be exercised: (1) the percolator used should be cylindrical or semicylindrical and cone-shaped as it nears the lower orifice; (2) a coarse granular sugar must be used, otherwise it will coalesce into a compact mass, which the liquid cannot permeate; (3) the purified cotton must be introduced with care.

If pressed in too tightly, the cotton will stop the process effectually; if inserted too loosely, the liquid will pass through the cotton rapidly and the filtrate will be weak and turbid (from imperfect filtration); it should be inserted completely within the neck of the percolator, since a protruding end, inside the percolator, up through the sucrose, will permit the last portions of water to pass out at the lower orifice without dissolving all the sucrose. For specific directions see *Syrups* (page 1301). The process of percolation is applied on a commercial scale for the making of official syrups as well as those for confectionary use.

Percolation is the preferred method for the preparation of Syrup USP (page 1301). The sucrose, in this instance, is placed in the percolator. However, a slightly modified approach must be used if a drug of vegetable origin is to be incorporated into the syrup. For example, wild cherry bark is first percolated with water; the collection vessel contains sucrose (800 g) and glycerol (50 mL). When the total volume is 1000 mL, the percolate is agitated to produce Wild Cherry Syrup PC.

Reconstitution—In order to improve stability and minimize microbial contamination, dry syrup formulations can be prepared and Purified Water USP added just prior to dispensing or use. Powder mixtures, wholly granulated products and partially granulated products have been investigated for this purpose by Ryder.[10]

The powder mixture preparation requires less equipment and energy to prepare. Chemical stability problems are minimal, since no heat or solvents are used in the process and a low moisture content can be obtained in the final product; unfortunately, powder mixtures are prone to homogeneity problems. In the case of the wholly granulated product all the ingredients are included in the granulation stage. The drug may be incorporated into the dry product before granulation or dissolved or suspended in the granulating fluid. After formation, the granules are dried and then screened to break down oversize particles. The advan-

tages of granulated over powder mixtures include better appearance, better flow, fewer segregation problems and less dust during processing. Partially granulated mixtures are used to gain some of the advantages of granulation without the disadvantages. Usually the drug, and other fine particles, are included at the granulation stage, perhaps with some diluents to improve flow and reduce segregation and dust. Materials selected for mixing with the dried granules would include thermolabile excipients, such as flavors, and free flowing materials, such as sugars.

Preservation—Syrups should be made in quantities which can be consumed within a few months, except in those cases where special facilities can be employed for their preservation; a low temperature is the best method. The USP indicates that syrups should not be exposed to excessive heat. Concentration without super-saturation is also a condition favorable to preservation. The USP states that syrups may contain preservatives to prevent bacterial and mold growth such as glycerin, methylparaben, benzoic acid and sodium benzoate, particularly when the concentration of sucrose in the syrup is low. Combinations of alkyl esters of *p*-hydroxybenzoic acid are effective inhibitors of yeasts which have been implicated in the contamination of commercial syrups. Any attempt to restore syrups spoiled through fermentation by heating them and "working them over" is reprehensible.

The official syrups should be preserved in well-dried bottles, preferably those which have been sterilized. These bottles should not hold more than is likely to be required during 4 to 6 weeks and should be filled completely, stoppered carefully and stored in a cool, dark place.

Syrups Prepared from Juices

Blackberry, pineapple and strawberry syrups may be prepared by following the directions in Raspberry Syrup PC. One volume of the concentrated raspberry juice is diluted with 11 volumes of syrup. Black Current Syrup PC is prepared in a similar manner but also can be prepared from black currants, with certain modifications. The pectin in the juice is destroyed with pectinase. The syrup is prepared from 700 g of sucrose and 560 mL of clarified juice and is preserved with sulfurous acid or sodium metabisulfite. The addition of a dye is permitted, provided it complies with the pertinent government regulations. Cherry Syrup USP is prepared from cherry juice by the addition of alcohol, sucrose and water (page 1301).

Honeys

Honeys are thick liquid preparations somewhat allied to the syrups, differing in that honey, instead of syrup, is used as a base. They are unimportant as a class of preparations today but at one time, before sugar was available and honey was the most common sweetening agent, they were used widely. PC lists two preparations containing honey. The first, Oxymel, or "acid honey," is a mixture of acetic acid (150 mL), purified water (150 mL) and honey (sufficient to produce 1000 mL of product). Squill Oxymel contains squill, water, acetic acid and honey and is prepared by a maceration process.

One nonofficial preparation contains borax (10.5 g), glycerin (5.25 g) and sufficient honey to make 1000 g. It has been indicated that this type of product can cause serious boric acid intoxication in babies. It should not be used in pharmaceutical practice. Thick and thin sugar pastes containing Caster sugar (very fine granular sugar), icing sugar (additive-free), polyethylene glycol 400 and hydrogen peroxide (in a final concentration of 0.15%) have been prepared and shown to be beneficial in the process of wound healing.

Mucilages

The official mucilages are thick, viscid, adhesive liquids, produced by dispersing gum in water, or by extracting the mucilaginous principles from vegetable substances with water. The mucilages all are prone to decomposition, showing appreciable decrease in viscosity on storage; they should never be made in quantities larger than can be used immediately, unless a preservative is added. Acacia Mucilage NF XII contains benzoic acid and Tragacanth Mucilage BPC (1973) contains alcohol and chloroform water. Chloroform in manufactured products for internal use is banned in some countries.

Acacia Mucilage may be prepared by placing 350 g of acacia in a graduated bottle, washing the drug with cold purified water, allowing it to drain and adding enough warm purified water, in which 2 g of benzoic acid has been dissolved, to make the product measure 1000 mL. The bottle then is stoppered, placed on its side, rotated occasionally and the product strained when the acacia has dissolved.

Tragacanth Mucilage BPC (1973) is prepared by mixing 12.5 g of tragacanth with 25 mL alcohol (90%) in a dry bottle and then quickly adding sufficient chloroform water to 1000 mL and shaking vigorously. The alcohol is used to disperse the gum to prevent agglomeration on addition of the water.

Mucilages are used primarily to aid in suspending insoluble subtances in liquids; their colloidal character and viscosity help prevent immediate sedimentation. Examples include sulfur in lotions, resin in mixtures and oils in emulsions. Both tragacanth and acacia either are partially or completely insoluble in alcohol. Tragacanth is precipitated from solution by alcohol, but acacia, on the other hand, is soluble in diluted alcoholic solutions. A 60% solution of acacia may be prepared with 20% alcohol and a 4% solution of acacia may be prepared even with 50% alcohol.

The viscosity of tragacanth mucilage is reduced by acid, alkali or sodium chloride, particularly if the mucilage is heated. It shows maximum viscosity at pH 5. Acacia is hydrolyzed by dilute mineral acids to arabinose, galactose, aldobionic and galacturonic acids. Its viscosity is low but is maintained over a wide pH range.

Recent research on mucilages includes the preparation of mucilage from plantain and the identification of its sugars, the preparation and suspending properties of cocoa gum, the preparation of glycerin ointments using flaxseed mucilage and the consideration of various gums and mucilages obtained from several Indian plants for pharmaceutical purposes.

Several synthetic mucilage-like substances such as *polyvinyl alcohol, methylcellulose, carboxymethylcellulose* and related substances, as described in Chapter 66, are used as mucilage substitutes, emulsifying and suspending agents. Methylcellulose (page 1306) is used widely as a bulk laxative since it absorbs water and swells to a hydrogel in the intestine, in much the same manner as *psyllium* or *karaya gum*. Methylcellulose Oral Solution is a flavored solution of the agent. It may be prepared by adding slowly the methylcellulose to about one-third the amount of boiling water, with stirring, until it is thoroughly wetted. Cold water then should be added and the wetted material allowed to dissolve while stirring. The viscosity of the solution will depend upon the concentration and the specifications of the methylcellulose. The synthetic gums are nonglycogenetic and may be used in the preparation of diabetic syrups. Several formulas for such syrups, based on sodium carboxymethylcellulose, have been proposed.

Uniformly smooth mucilages sometimes are difficult to prepare due to the uneven wetting of the gums. In general, it is best to use fine gum particles and disperse them with good agitation in a little 95% alcohol or in cold water (except for methylcellulose). The appropriate amount of water then can be added with constant stirring. A review of the chemistry and properties of acacia and other gums has been prepared.[1]

Jellies

Jellies are a class of gels in which the structural coherent matrix contains a high portion of liquid, usually water. They are similar to mucilages, in that they may be prepared from similar gums, but they differ from the latter in having a jelly-like consistency. A whole gum of the best quality, rather than a powdered gum, is desirable in order to obtain a clear preparation of uniform consistency. Tragacanth is the gum used in the preparation of Ephedrine Sulfate Jelly NF XII. While the specific thickening agent in the USP jellies is not indicated, reference usually is made in the monograph to a water-soluble viscous base. These preparations also may be formulated with water from acacia, chondrus, gelatin, carboxymethylcellulose and similar substances.

Jellies are used as lubricants for surgical gloves, catheters and rectal thermometers. Lidocaine Hydrochloride Jelly USP is used as a topical anesthetic. Therapeutic vaginal jellies are available and certain jelly-like preparations are used for contraceptive purposes, which often contain surface-active agents to enhance the spermatocidal properties of the jelly. Aromatics, such as methyl salicylate and eucalyptol, often are added to give the preparation a desirable odor.

Jellies are prone to microbial contamination and therefore contain preservatives, eg, methyl *p*-hydroxybenzoate is used as a preservative in a base for medicated jellies. This base contains sodium alginate, glycerin, calcium gluconate and water. The calcium ions cause a cross-linking with sodium alginate to form a gel of firmer consistency. A discussion of gels is provided later in the chapter.

Nonaqueous Solutions

It is difficult to evaluate fairly the importance of nonaqueous solvents in pharmaceutical processes. That they are important in the manufacture of pharmaceuticals is an understatment. However, pharmaceutical preparations, and, in particular, those intended for internal use, rarely contain more than minor quantities of the organic solvents that are common to the manufacturing or analytical operation. For example, industry uses large quantities of chloroform in some operations but the solvent is of only minor importance with respect to the final product. One mL of chloroform dissolves in about 200 mL of water and the solution so formed finds some use as a vehicle (see the section on *Aromatic Waters*). Chloroform has been an ingredient in a number of cough syrups in the past but it has been banned in the US by the FDA in manufactured products intended for internal use. Solvents such as acetone, benzene and petroleum ether must not be ingredients in preparations intended for internal use.

Products of commerce may contain solvents such as ethanol, glycerin, propylene glycol, certain oils and liquid paraffin. Preparations intended for external use may contain ethanol, methanol, isopropyl alcohol, polyethylene glycols, various ethers and certain esters. A good example of preparations of this type are the rubefacient rubbing alcohols. Rubbing Alcohol must be manufactured in accordance with the requirements of the Bureau of Alcohol, Tobacco and Firearms, US Treasury Dept, using Formula 23-H denatured alcohol. This mixture contains 8 parts by volume of acetone, 1.5 parts by volume of methyl isobutyl ketone and 100 parts by volume of ethanol. Besides the alcohol in the Rubbing Alcohol, the final product must contain water, sucrose octaacetate or denatonium benzoate and may contain color additives, perfume oils and a suitable stabilizer. The alcohol content, by volume, is not less than 68.5% and not more than 71.5%. The isopropyl alcohol content in Isopropyl Rubbing Alcohol can vary from 68.0% to 72.0% and the finished product may contain color additives, perfume oils and suitable stabilizers.

Although the lines between aqueous and nonaqueous preparations tend to blur in those cases where the solvent is water-soluble, it is possible to categorize a number of products as nonaqueous. This section is, therefore, devoted to groups of nonaqueous solutions; the alcoholic or hydroalcoholic solutions (eg, elixirs and spirits), ethereal solutions (eg, collodions), glycerin solutions (eg, glycerins), oleaginous solutions (eg, liniments, oleovitamins and toothache drops), inhalations and inhalants.

Although this list is self-limiting, a wide variety of solvents are used in various pharmaceutical preparations. Solvents such as glycerol formal, dimethylacetamide and glycerol dimethylketal have been recommended for many products produced by the industry. However, the toxicity of many of these solvents is not well-established and, for this reason, careful clinical studies should be carried out on the formulated product before it is released to the marketplace.

It is essential that the toxicity of solvents be tested appropriately and approved in order to avoid problems: for example, the tragic loss of life which occurred during 1937 when diethylene glycol was used in an elixir of sulfanilamide. The result of this tragedy was the 1938 Federal Food, Drug and Cosmetic Act, which required that products be tested for both safety and effectiveness.

Collodions

Collodions are liquid preparations containing pyroxylin (a nitrocellulose) in a mixture of ethyl ether and ethanol. They are applied to the skin by means of a soft brush or other suitable applicator and, when the ether and ethanol have evaporated, leave a film of pyroxylin on the surface. The official medicated collodion, Salicylic Acid Collodion USP, contains 10% *w/v* of salicylic acid in Flexible Collodion USP and is used as a keratolytic agent in the treatment of corns and warts. Collodion USP and Flexible Collodion USP are water-repellent protectives for minor cuts and scratches. Collodion is made flexible by the addition of castor oil and camphor. Collodion has been used to reduce or eliminate the side effects of fluorouracil treatment of solar keratoses. Vehicles other than flexible collodion, such as a polyacrylic base, have been used to incorporate salicylic acid for the treatment of warts with less irritation.

Elixirs

Elixirs are clear, pleasantly flavored, sweetened hydroalcoholic liquids intended for oral use. They are used as flavors and vehicles such as Aromatic Elixir (page 1302) for drug substances and, when such substances are incorporated into the specified solvents, they are classified as medicated elixirs, eg, Dexamethasone Elixir USP and Phenobarbital Elixir USP. The main ingredients in elixirs are ethanol and water but glycerin, sorbitol, propylene glycol, flavoring agents, preservatives and syrups often are used in the preparation of the final product.

The distinction between some of the medicated syrups and elixirs is not always clear. For example, Ephedrine Sulfate Syrup USP contains between 20 and 40 mL of alcohol in 1000 mL of product. Ephedrine Elixir PC contains syrup and 100 mL of ethanol in the same final volume. Definitions are, therefore, inconsistent and, in some in-

stances, not too important with respect to the naming of the articles of commerce. The exact composition must, however, be known if the presence or absence of an ingredient (eg, sucrose) is of therapeutic significance or when an additional ingredient must be incorporated in the product.

Elixirs contain ethyl alcohol. However, the alcoholic content will vary greatly, from elixirs containing only a small quantity to those that contain a considerable portion as a necessary aid to solubility. For example, Aromatic Elixir USP contains 21 to 23% of alcohol; Compound Benzaldehyde Elixir, on the other hand, contains 3 to 5%.

Elixirs also may contain glycerin and syrup. These may be added to increase the solubility of the medicinal agent or for sweetening purposes. Some elixirs contain propylene glycol. Claims have been made for this solvent as a satisfactory substitute for both glycerin and alcohol. Sumner,[12] in his paper on terpin hydrate preparations, summarized the advantages and disadvantages of this solvent and suggested several formulations with therapeutic characteristics superior to those of the elixir described in NF XIII.

One usual dose of the elixir (5 mL) contains 85 mg of terpin hydrate. This substance is used in bronchitis in doses of 125 to 300 mg as an expectorant. Therefore, the elixir is ineffective for the treatment of bronchitis. However, it is used as a vehicle for the drugs in many commercially available cough syrups. These may contain dextromethorphan hydrobromide codeine phosphate, chlorpheniramine maleate, pyrilamine maleate, ammonium chloride, creosote and a wide variety of other drugs with expectorant and antitussive properties.

One of the four formulations described in Sumner's paper is given below:

Terpin Hydrate	6.0 g
Orange Oil	0.1 mL
Benzaldehyde	0.005 mL
Sorbitol Solution USP	10.0 mL
Propylene Glycol	40.0 mL
Alcohol	43.0 mL
Purified Water, a sufficient quantity, to make	100.0 mL

Dissolve the terpin hydrate in the propylene glycol and sorbitol solution which have been heated to 50°. Add the oil and the benzaldehyde to the alcohol and mix with the terpin hydrate solution at 25°. Add sufficient purified water to make the product measure 100 mL.

The elixir contains 300 mg of terpin hydrate/5 mL, a minimal quantity of alcohol and flavoring agents which adequately mask the taste of propylene glycol.

Although alcohol is an excellent solvent for some drugs, it does accentuate the saline taste of bromides and similar salts. It often is desirable, therefore, to substitute some other solvent that is more effective in masking such tastes for part of the alcohol in the formula. In general, if taste is a consideration, the formulator is more prone to use a syrup rather than a hydroalcoholic vehicle.

An elixir may contain water- and alcohol-soluble ingredients. If such is the case, the following procedure is indicated:

Dissolve the water-soluble ingredients in part of the water. Add and solubilize the sucrose in the aqueous solution. Prepare an alcoholic solution containing the other ingredients. Add the aqueous phase to the alcoholic solution, filter and make to volume with water.

Sucrose increases viscosity and decreases the solubilizing properties of water and so must be added after primary solution has been effected. A high alcoholic content is maintained during preparation by adding the aqueous phase to the alcoholic solution. Elixirs always should be brilliantly clear. They may be strained or filtered and, if necessary, subjected to the clarifying action of purified talc or siliceous earth.

One of the former official elixirs, Iso-Alcoholic Elixir NF XV (page 1328), actually is a combination of two solutions, one containing 8 to 10% ethanol and the other containing 73 to 78%. It is used as a vehicle for various medicaments that require solvents of different alcoholic strengths. For example, the alcoholic strength of the elixir to be used with a single liquid galenical is approximately the same as that of the galenical. When different alcoholic strengths are employed in the same prescription, the elixir to be used is the one that produces the best solution. This is usually the average of the alcoholic strengths of the several ingredients. For nonextractive substances, the lowest alcoholic strength of elixir that will produce a clear solution should be selected.

The formula for High-Alcoholic Elixir is:

Compound Orange Spirit	4 mL
Saccharin	3 g
Glycerin	200 mL
Alcohol, a sufficient quantity, to make	1000 mL

This elixir, and many other liquid preparations intended for internal use (eg, the diabetic syrups thickened with sodium carboxymethylcellulose or similar substances) contain saccharin. During the past few years, scientists have studied the toxic effects of this sweetening agent and of the cyclamates. The cyclamate studies showed that the sweetener could produce cancer in animals and, as a result, this substance was removed from a wide variety of products. Similar studies have been carried out on saccharin.

Cyclamates and saccharin have been banned in some countries as ingredients in manufactured products. Much reserch has been done to find a safe synthetic substitute for sucrose. As a result, aspartame (methyl *N*(-L-α-asparty1)-L-phenylalaninate), which is about 200 times sweeter than sucrose, is being used now in many commercial preparations as the sweetening agent. It is sparingly soluble in water and is most stable at a pH of 4.3. This compound likely will be used in a number of pharmaceutical formulations in the future.[13]

Incompatibilities—Since elixirs contain alcohol, incompatibilities of this solvent are an important consideration during formulation. Alcohol precipitates tragacanth, acacia and agar from aqueous solutions. Similarly, it will precipitate many inorganic salts from similar solutions. The implication here is that such substances should be absent from the aqueous phase or present in such concentrations that there is no danger of precipitation on standing.

If an aqeous solution is added to an elixir, a partial precipitation of ingredients may occur. This is due to the reduced alcoholic content of the final preparation. Usually, however, the alcoholic content of the mixture is not sufficiently decreased to cause separation. As vehicles for tinctures and fluidextracts, the elixirs generally cause a separation of extractive matter from these products due to a reduction of the alcoholic content.

Many of the incompatibilities between elixirs, and the substances combined with them, are due to the chemical characteristics of the elixir *per se*, or of the ingredients in the final preparation. Thus, certain elixirs are acid in reaction while others may be alkaline and will, therefore, behave accordingly.

Glycerins

Glycerins or glycerites are solutions or mixtures of medicinal substances in not less than 50% by weight of glycerin. Most of the glycerins are extremely viscous and some are of a jelly-like consistency. Few of them are used extensively.

Glycerin is a valuable pharmaceutical solvent forming permanent and concentrated solutions not otherwise obtainable. Some of these solutions are used in their original form as medicinal agents while others are used to prepare aqueous and alcoholic dilutions of substances which are not readily soluble in water or alcohol. Antipyrine and Benzocaine Otic Solution USP was discussed previously under *Otic Solutions*. One of the glycerins, Phenol Glycerin PC is diluted with glycerin to form the pharmaceutical preparation, Phenol Ear-Drops PC.

Phenol Glycerin PC

Phenol	**160 g**
Glycerin	**840 g**

Dissolve the phenol in the glycerin.

Phenol Ear-Drops PC

Phenol Glycerin	**40 mL**
Glycerin, a sufficient quantity, to make	**100 mL**

Water must not be added to this preparation. It reacts with the phenol to produce a preparation which is caustic and, consequently, damaging to the area of application. This product no longer is recommended because of the possibility of necrosis and perforation of the tympanic membrane. As noted under *Otic Solutions*, glycerin alone is used to aid in the removal of cerumen.

Sodium Bicarbonate Ear-Drops PC may be used if wax is to be removed from the ear. This preparation contains sodium bicarbonate (5 g), glycerin (30 mL) and purified water (a sufficient quantity to make 100 mL). A glycerin base was chosen as the optimum solvent for an otic preparation in a study involving the stability and antimicrobial activity of kanamycin sulfate otic drops.

Starch Glycerin, an emollient, contains starch (100 g), benzoic acid (2 g), purified water (200 mL) and glycerin (700 mL).

Glycerins are hygroscopic and should be stored in tightly closed containers.

Inhalations and Inhalants

Inhalations

These preparations are so used or designed that the drug is carried into the respiratory tree of the patient. The vapor or mist reaches the affected area and gives prompt relief from the symptoms of bronchial and nasal congestion. The USP defines Inhalations in the following way:

> Inhalations are drugs or solutions of drugs administered by the nasal or oral respiratory route for local or systemic effect. Examples in this Pharmacopeia are Epinephrine Inhalation and Isoproterenol Hydrochloride Inhalation. Nebulizers are suitable for the administration of inhalation solutions only if they give droplets sufficiently fine and uniform in size so that the mist reaches the bronchioles.
>
> Another group of products, also known as inhalations, and sometimes called insufflations, consists of finely powdered or liquid drugs that are carried into the respiratory passages by the use of special delivery systems, such as pharmaceutical aerosols, that hold a solution or suspension of the drug in a liquefied gas propellant (see Aerosols). When released through a suitable valve and oral adapter, a metered dose of the inhalation is propelled into the respiratory tract of the patient. Powders also may be administered by mechanical devices that require a manually produced pressure or a deep inspiration by the patient, eg, *Cromolyn Sodium*.
>
> Solutions may be nebulized by use of inert gases. Nebulized solutions may be breathed directly from the nebulizer, or the nebulizer may be attached to a plastic face mask, tent or intermittent positive-pressure breathing (IPPB) machine.

As stated in the USP, particle size is of major importance in the administration of this type of preparation. The various mechanical devices that are used in conjunction with inhalations are described in some detail in Chapter 104. It has been reported that the optimum particle size for penetration into the pulmonary cavity is of the order of 0.5 to 7 μm. Fine mists are produced by pressurized aerosols and hence possess basic advantages over the older nebulizers; in addition, metered aerosols deliver more uniform doses. See Chapter 92.

The term *Inhalation* is used commonly by the layman to represent preparations intended to be vaporized with the aid of heat, usually steam, and inhaled. Benzoin Inhalation PC contains benzoin, storax and alcohol. The vapors from a preparation containing 1 teaspoonful of the tincture and 1 qt of boiling water may be inhaled. The device known as a *vaporizer* is used with a number of commercially available preparations of this type.

Epinephrine Inhalation and Isoproterenol Hydrochloride Inhalation are described in the USP.

Inhalants

The USP defines inhalants as follows:

> A special class of inhalations termed "inhalants" consists of drugs or combinations of drugs that, by virtue of their high vapor pressure, can be carried by an air current into the nasal passage where they exert their effect. The container from which the inhalant is administered is known as an inhaler.

Propylhexedrine Inhalant and Tuaminoheptane Inhalant consist of cylindrical rolls of suitable fibrous material impregnated with propylhexedrine or tuaminoheptane (as carbonate), usually aromatized, and contained in a suitable inhaler. Propylhexedrine is the active ingredient in the widely used Benzedrex Inhaler. Both of these drugs are vasoconstrictors used to relieve nasal congestion. Inhalers which come in contact with the mouth or nasal passages become contaminated by bacteria, thus, they should be restricted to personal use.

Another inhalant is Amyl Nitrite USP which is very flammable and should not be used where it may be ignited. It is packaged in sealed glass vials in a protective gauze. Upon breaking the vial, the gauze absorbs the drug which is then inhaled for the treatment of anginal pain. See page 843.

Liniments

Liniments are solutions or mixtures of various substances in oil, alcoholic solutions of soap or emulsions. They are intended for external application and should be so labeled. They are rubbed onto the affected area and, because of this, were once called *embrocations*. Dental liniments, which are no longer official, are solutions of active substances and are rubbed into the gums. Most dentists question their usefulness and, consequently, this type of preparation is relatively unimportant as a pharmaceutical form.

Liniments usually are applied with friction and rubbing of the skin, the oil or soap base providing for ease of application and massage. Alcoholic liniments are used generally for their rubefacient, counterirritant, mildly astringent and penetrating effects. Such liniments penetrate the skin more readily than do those with an oil base. The oily liniments, therefore, are milder in their action but are more useful when massage is required. Depending on their ingredients, such liniments may function solely as protective coatings. Liniments should not be applied to skin that is bruised or broken.

Many of the marketed "white" liniments are based on the formulation below or variations thereof.

White Liniment PC

Ammonium Chloride	**12.5 g**
Dilute Ammonia Solution	**45 mL**
Oleic Acid	**85 mL**

Turpentine Oil	250	mL
Water for Preparations	625	mL

Mix the oleic acid with the turpentine oil. Add the dilute ammonia solution mixed with 4h mL of previously warmed water and shake. Dissolve the ammonium chloride in the remainder of the water, add to the emulsion and mix.

Other liniments contain antipruritics, astringents, emollients or analgesics and are classified on the basis of their active ingredient. An example is:

Compound Calamine Application PC
(Compound Calamine Liniment)

Calamine	100 g
Zinc Oxide	50 g
Wool Fat	25 g
Zinc Stearate	25 g
Yellow Soft Paraffin	250 g
Liquid Paraffin	550 g

The powders are triturated to a smooth paste with some of the liquid paraffin (Liquid Petrolatum). The wool fat, zinc stearate and yellow soft paraffin (Petrolatum) are melted, mixed with some of the liquid paraffin, the mixture incorporated with the triturated powders and the rest of the liquid paraffin added with mixing.

Dermatologists prescribe products of this type but only those containing the rubefacients are advertised extensively and used by consumers for treating minor muscular aches and pains.

Because of the confusion of camphorated oil (camphor liniment) with castor oil, which has resulted in ingestion and, perhaps, to poisoning, camphorated oil has been banned from the market. It is essential that these applications be marked clearly for external use only. (Camphorated Oil presently is classified as a new drug by the FDA.)

Oleovitamins

Oleovitamins are fish-liver oils diluted with edible vegetable oil or solutions of the indicated vitamins or vitamin concentrates (usually vitamin A and D) in fish-liver oil. The definition is broad enough to include a wide variety of marketed products.

Oleovitamin A and D is official; vitamin D may be present as ergocalciferol or cholecalciferol obtained by the activation of ergosterol or 7-dehydrocholesterol or may be obtained from natural sources. Synthetic vitamin A, or a concentrate, may be used to prepare oleovitamin A. The starting material for the concentrate is a fish-liver oil, the active ingredient being isolated by molecular distillation or by a saponification and extraction procedure. The latter procedure is described in detail in the monograph for Concentrated Vitamin A Solution PC.

These vitamins are unstable in the presence of rancid oils and, therefore, these preparations and, in particular, Oleovitamin A, should be stored in small, tight containers, preferably under vacuum or under an atmosphere of an inert gas, protected from light.

Spirits

Spirits, popularly known as essences, are alcoholic or hydroalcoholic solutions of volatile substances. Like the aromatic waters, the active ingredient in the spirit may be a solid, liquid or gas. The genealogical tree for this class of preparations begins with the distinguished pair of products, Brandy (*Spiritus Vini Vitis*) and Whisky (*Spiritis Frumenti*), and ends with a wide variety of products that comply with the definition given above. Physicians have debated the therapeutic value of the former products and these are no longer official in the compendia.

Some of these spirits are used internally for their medicinal value, a few medicinally by inhalation and a large number as flavoring agents. The latter group provides a convenient and ready means of obtaining the volatile oil in the proper quantity. For example, a spirit or spirit-like preparation may be used in the formulation of aromatic waters or other pharmaceuticals that require a distinctive flavor.

Spirits should be stored in tight, light-resistant containers and in a cool place. This prevents evaporation and volatilization of either the alcohol or the active principle.

Preparation—There are four classic methods of preparation:

Simple Solution—This is the method by which the majority of spirits are prepared. The formula and procedure given for Aromatic Ammonia Spirit USP illustrate this method of preparation.

Aromatic Ammonia Spirit USP

Ammonium Carbonate, in translucent pieces	34 g
Strong Ammonia Solution	36 mL
Lemon Oil	10 mL
Lavender Oil	1 mL
Nutmeg Oil	1 mL
Alcohol	700 mL
Purified Water, a sufficient quantity to make	1000 mL

Dissolve the ammonium carbonate in the strong ammonia solution and 195 mL of purified water by gentle agitation and allow the solution to stand for 12 hours. Dissolve the oils in the alcohol, contained in a graduated bottle or cylinder, and gradually add the ammonium carbonate solution and enough purified water to make the product measure 1000 mL. Set the mixture aside in a cool place for 24 hours, occasionally agitating it, then filter, using a covered funnel.

The spirit is a respiratory stimulant and is administered by inhalation of the vapor as required. It is marketed in suitable tight, light-resistant containers but is also available in a single-dose glass vial wrapped in a soft cotton envelope. The vial is broken easily; the cotton acts as a sponge for the spirit.

Ammonium carbonate is a mixture of ammonium bicarbonate and ammonium carbamate (NH_2COONH_4). The carbamate reacts with water to form the carbonate. An ammonium carbonate solution is, therefore, a solution of ammonium bicarbonate and ammonium carbonate in water. However, it decomposes in water, the decomposition products being ammonia, carbon dioxide and water. The stability of the spirit is improved by the addition of strong ammonia solution. This represses the hydrolysis of ammonium carbonate and, in this way, decreases the loss of dissolved gases.

Solution with Maceration—In this procedure, the leaves of a drug are macerated in purified water to extract water-soluble matter. They are expressed and the moist, macerated leaves are added to a prescribed quantity of alcohol. The volatile oil is added to the filtered liquid. Peppermint Spirit USP is made by this process. Peppermint Spirit PC differs from the official product in that it is a solution of the volatile oil in alcohol only. The concentration of volatile oil in the final product is about the same but the official preparation possesses a green color. The ready availability of soluble chlorophyll and other coloring agents had led to the frequent suggestion that a more uniform product could be obtained through their use. However, these agents cannot be used in preparing the official article.

The formula and procedure for Peppermint Spirit USP (page 798) illustrate this method of preparation.

Chemical Reaction—No official spirits are prepared by this process. Ethyl nitrite is made by the action of sodium nitrite on a mixture of alcohol and sulfuric acid in the cold. This substance then is used to prepare Ethyl Nitrite Spirit, a product no longer official.

Distillation—Brandy and Whisky are made by distillation. The latter is derived from the fermented mash of wholly or partially germinated malted cereal grains and the former from the fermented juice of ripe grapes.

Incompatibilities—Spirits are, for the most part, preparations of high alcoholic strength and do not lend themselves well to dilution with aqueous solutions or liquids of low alcoholic content. The addition of such a solution invariably causes separation of some of the material dissolved in the spirit, evidenced by a turbidity which, in time, may disappear as distinct layering occurs. Salts may be precipi-

tated from their aqueous solutions by the addition of spirits due to their lesser solubility in alcoholic liquids.

Some spirits show incompatibilities characteristic of the ingredients they contain. For example, Aromatic Ammonia Spirit cannot be mixed with aqueous preparations containing alkaloids (eg, codeine phosphate). An acid-base reaction (ammonia-phosphate) occurs and, if the alcohol content of the final mixture is too low, codeine will precipitate.

Toothache Drops

Toothache drops are preparations used for temporary relief of toothache by application of a small pledget of cotton saturated with the product into the tooth cavity. Anesthetic compounds include clove oil, eugenol or benzocaine; other ingredients include camphor, creosote, menthol and alcohol.

These preparations no longer are recognized officially. Furthermore, dentists do not recommend the use of toothache drops if the patient has ready access to adequate dental services. The preparations may damage the gums and produce complications more severe than the original toothache. However, many areas do not have adequate dental services and the pharmacist will, of necessity, handle these preparations, and he should warn the patient of possible hazards associated with their use.

Toothache Drops NF XI contains 25 g of chlorobutanol in sufficient clove oil to make the product measure 100 mL. Another formulation contains creosote, clove oil, benzocaine and alcohol in a flexible collodion base.

Emulsions

An emulsion is a two-phase system prepared by combining two immiscible liquids, one of which is dispersed uniformly throughout the other and consists of globules that have diameters equal to or greater than those of the largest colloidal particles. The globule size is critical, of course, and must be such that the system achieves maximum stability. However, even under the best conditions, separation of the two phases will occur unless a third substance, an *emulsifying agent*, is incorporated. The basic emulsion must, therefore, contain three components, but the products of commerce may consist of a number of therapeutic agents dissolved in either of the two phases.

Most emulsions incorporate an aqueous phase into a nonaqueous phase (or *vice versa*). However, it is possible to prepare emulsions that are basically nonaqueous. For example, investigations of the emulsifying effects of anionic and cationic surfactants on the nonaqueous immiscible system, glycerin and olive oil, have shown that certain amines and three cationic agents produced stable emulsions. This broadening of the basic definition for the term *emulsion* is recognized in the USP.

An emulsion is a two-phase system in which one liquid is dispersed in the form of small droplets throughout another liquid. The dispersed liquid is known as the internal or discontinuous phase, whereas the dispersion medium is known as the external or continuous phase. Where oil is the dispersed phase and an aqueous solution is the continuous phase, the system is designated as an oil-in-water (O/W) emulsion and can be diluted easily and uniformly with water. Conversely, where water, or an aqueous solution is the dispersed phase, and oil, or oleaginous material, is the continuous phase, the system is designated as a water-in-oil (W/O) emulsion.

Many emulsifying agents (or emulsifiers) are available, among them the following:

Natural Emulsifying Agents—These substances may be derived from either animal or vegetable sources. Examples of those obtained from the former source are gelatin, egg yolk, casein, wool fat or cholesterol. Acacia, tragacanth, chondrus or pectin are representative of those obtained from vegetable sources. Various cellulose derivatives, eg, methylcellulose and carboxymethylcellulose, are used to increase the viscosity of the aqueous phase and thereby enhance emulsion stability.

Finely Divided Solids—Examples are bentonite, magnesium hydroxide, aluminum hydroxide or magnesium trisilicate.

Synthetic Emulsifying Agents—This group may be subdivided further into the anionic, cationic or nonionic agents. Examples are, in order of presentation, sodium lauryl sulfate, benzalkonium chloride or polyethylene glycol 400 monostearate.

Many of these emulsifying agents are described in greater detail in Chapter 66.

In NF XIII it was suggested that only O/W emulsions are suitable for oral use because these are water-miscible and thus their oiliness is masked. This compendium gave specific directions for the preparation of emulsions using gelatin as an emulsifying agent. These preparations are based on either type A or type B gelatin.

Type A gelatin is prepared by acid-treated precursors and is used at a pH of about 3.2. It is incompatible with anionic emulsifying agents such as the vegetable gums. The following formula was recommended:

Gelatin (Type A)	6 g
Tartaric Acid	0.6 g
Flavor as desired	
Alcohol	60 mL
Oil	500 mL
Purified Water, to make	1000 mL

Add the gelatin and the tartaric acid to about 300 mL of purified water, allow to stand for a few minutes, heat until the gelatin is dissolved, then raise the temperature to about 98° and maintain this temperature for about 20 min. Cool to 50°, add the flavor, the alcohol and sufficient purified water to make 500 mL. Add the oil, agitate the mixture thoroughly and pass it through a homogenizer or a colloid mill until the oil is dispersed completely and uniformly.

This emulsion cannot be prepared by trituration or by the use of the usual stirring devices.

Type B gelatin is prepared from alkali-treated precursors and is used at a pH of about 8.0. It may be used with other anionic emulsifying agents but is incompatible with cationic types. If the emulsion contains 50% oil, 5 g of Type B gelatin, 2.5 g of sodium bicarbonate and sufficient tragacanth or agar should be incorporated into the aqueous phase to yield 1000 mL of product of the required viscosity.

The emulsion type (O/W or W/O) is of lesser significance if the final preparation is to be applied to the skin. If there are no breaks in the skin, a W/O emulsion can be applied more evenly since the skin is covered with a thin film of sebum. The latter substance favors the oily phase and contributes to the ease of application. The choice of emulsion type will, however, depend on many other factors. This particularly is true for those preparations which have basic cosmetic characteristics. It may be advantageous to formulate an O/W emulsion if ease of removal is an important consideration to the patient.

An emulsion that may be prepared by the mortar and pestle method is the following Mineral Oil Emulsion USP.

Mineral Oil	500 mL
Acacia, in very fine powder	125 g
Syrup	100 mL
Vanillin	40 mg
Alcohol	60 mL
Purified Water, a sufficient quantity	1000 mL

The mineral oil and acacia are mixed in a dry Wedgwood mortar. Water (250 mL) is added and the mixture triturated vigorously until an emulsion is formed. A mixture of the syrup, 50 mL of purified water and the vanillin dissolved in alcohol is added in divided portions with trituration; sufficient purified water is then added to the proper volume, the mixture mixed well and homogenized.

Very few emulsions are included now in the official compendia. The PC suggests that the term "emulsion" be restricted to preparations, usually O/W, intended for internal use and contains the following: Liquid Paraffin Emulsion, Liquid Paraffin and Magnesium Hydroxide Emulsion, Liquid Paraffin and Phenolphthalein Emulsion and Concentrated Peppermint Emulsion.

This, however, should not lead the reader to the conclusion that emulsions are a relatively unimportant class of pharmaceuticals. While it is true that few preparations carry the term *emulsion* in their titles, they are of great significance as bases for other types of preparations, particularly in the dermatological and cosmetic areas. Academically, they illustrate the importance of the relationship between the theory and practice of emulsion technology and, practically, they possess a number of important advantages over other liquid forms. These may be summarized in the following way:

1. In an emulsion, the therapeutic properties and the spreading ability of the constituents are increased.
2. The unpleasant taste or odor of an oil can be masked partially or wholly, by emulsification. Secondary masking techniques are available to the formulator but these must be used with caution. If flavors and sweetening agents are added to the emulsion, only minimal amounts should be used in order to prevent the nausea or gastric distress that results on ingestion of larger quantities of these.
3. The absorption and penetration of medicaments are controlled more easily if they are incorporated into an emulsion.
4. Emulsion action is prolonged and the emollient effect is greater than that observed with comparable preparations.
5. Water is an inexpensive diluent and a good solvent for the many drugs and flavors that are incorporated into an emulsion.

The effects of viscosity, surface tension, solubility, particle size, complexation and excipients on the bioavailability of oral suspensions and emulsions have been discussed in detail by Rettig.[14]

The aqueous phase of the emulsion favors the growth of microorganisms and, because of this, a preservative usually is added to the product. Some of the preservatives that have been used include chlorocresol, chlorobutanol, mercurial preparations, salicylic acid, the esters of *p*-hydroxybenzoic acid, benzoic acid, sodium benzoate or sorbic acid. The preservative should be selected with regard for the ultimate use of the preparation and possible incompatibilities between the preservative and the ingredients in the emulsion, eg, binding between the surface-active agent and the preservative. Low pH values of 5 to 6 and low concentrations of water are characteristics also likely to inhibit microbiological growth in emulsions.

Most emulsions consist of a nonaqueous (or oil or lipid) phase and an aqueous (or water) phase, thus some of the preservative may pass into the oil phase and be removed from the aqueous phase. It is in the aqueous phase that microorganisms tend to grow. As a result, water-soluble preservatives are more effective since the concentration of the unbound preservative in the aqueous phase assumes a great deal of importance in inhibiting the microbial growth. Esters of *p*-hydroxybenzoic acid appear to be the most satisfactory preservatives for emulsions. Many mathematical models have been used to determine the availability of preservatives in emulsified systems. However, because of the number of factors which reduce the effectiveness of the preservative, a final microbiological evaluation of the emulsion should be performed.

While emphasis concerning preservation of emulsions deals with the aqueous phase, microorganisms can reside also in the lipid phase. Consequently, it has been recommended that pairs of preservatives be used to ensure adequate concentration in both phases. Esters of *p*-hydroxybenzoic acid can be used to ensure appropriate concentrations in both phases because of their difference in oil and water solubilities.

An emulsion can be diluted with the liquid that constitutes, or is miscible with, the external phase. The diluting liquid, however, will decrease the viscosity of the preparation and, in certain instances, invert the emulsion. The latter phenomena may occur if the emulsifier-in-water method (see below) is used to prepare the emulsion.

Preparation

The theory of emulsion preparation is discussed in Chapter 19. The following procedures are those suggested by Griffin *et al.*[15]

The formulator must first determine the physical and chemical characteristics of the active ingredient. He must know the following:

1. Structural formula
2. Melting point
3. Solubility
4. Stability
5. Dose
6. Specific chemical incompatibilities

It also is necessary, at this stage, to decide on the type of emulsion required. Washable emulsions are of the O/W type; nonwashable, the W/O type. In general, O/W emulsions contain over 70% water. W/O emulsions usually will contain higher concentrations of oils and waxes. The preparation of cream and ointment emulsions for topical use is given in Chapter 87.

Experimental formulations may be prepared by the following procedure:

1. Group the ingredients on the basis of their solubilities in the aqueous and nonaqueous phases.
2. Determine the type of emulsion required and calculate an approximate HLB (hydrophile-lipophile balance) value.
3. Blend a low HLB emulsifier and a high HLB emulsifier to the calculated value. For experimental formulations, use a higher concentration of emulsifier (eg, 10 to 30% of the oil phase) than that required to produce a satisfactory product. Emulsifiers should, in general, be stable chemically, nontoxic and suitably low in color, odor and taste. The emulsifier is selected on the basis of these characteristics, the type of equipment being used to blend the ingredients and the stability characteristics of the final product. Emulsions should not coalesce at room temperature, when frozen and thawed repeatedly or at elevated temperatures of up to 50°. Mechanical energy input varies with the type of equipment used to prepare the emulsion. The more the energy input, the less the demand on the emulsifier. Both process and formulation variables can affect the stability of an emulsion.
4. Dissolve the oil-soluble ingredients and the emulsifiers in the oil. Heat, if necessary, to approximately 5° to 10° over the melting point of the highest melting ingredient or to a maximum temperature of 70° to 80°.
5. Dissolve the water-soluble ingredients (except acids and salts) in a sufficient quantity of water.
6. Heat the aqueous phase to a temperature which is 3° to 5° higher than that of the oil phase.
7. Add the aqueous phase to the oily phase with suitable agitation.
8. If acids or salts are employed, dissolve them in water and add the solution to the cold emulsion.
9. Examine the emulsion and make adjustments in the formulation if the product is unstable. It may be necessary to add more emulsifier, to change to an emulsifier with a slightly higher or lower HLB value or to use an emulsifier with different chemical characteristics.

The technique of emulsification of pharmaceutical preparations has been described by White.[16] The preparation of an emulsion requires work to reduce the internal phase into small droplets and disperse them through the external phase. This can be accomplished by a mortar and pestle or a high-speed emulsifier. The addition of emulsifying agents not only reduces this work but also stabilizes the final emulsion. Emulsions may be prepared by four principle methods.

Addition of Internal Phase to External Phase—This is usually the most satisfactory method for preparing emulsions since there is always an excess of the external phase present which promotes the type of emulsion desired. If the external phase is water and the internal phase is oil, the water-soluble substances are dissolved in the water and the oil-

soluble substances mixed thoroughly in the oil. The oil mixture is added in portions to the aqueous preparation with agitation. Sometimes, in order to give a better shearing action during the preparation, all of the water is not mixed with the emulsifying agent until the primary emulsion with the oil is formed; subsequently, the remainder of the water is added. An example using gelatin Type A is given above.

Addition of the External Phase to the Internal Phase—Using an O/W emulsion as an example, the addition of the water (external phase) to the oil (internal phase) will promote the formation of a W/O emulsion due to the preponderance of the oil phase. After further addition of the water, phase inversion to an O/W emulsion should take place. This method especially is useful and successful when hydrophilic agents such as acacia, tragacanth or methylcellulose are first mixed with the oil, effecting dispersion without wetting. Water is added and, eventually, an O/W emulsion is formed. This "dry gum" technique is a rapid method for preparing small quantities of emulsion. The ratio 4 parts of oil, 2 parts of water and 1 part of gum provides maximum shearing action on the oil globules in the mortar. The emulsion then can be diluted and triturated with water to the appropriate concentrations. The preparation of Mineral Oil Emulsion described above is an example.

Mixing Both Phases after Warming Each—This method is used when waxes or other substances which require melting are used. The oil-soluble emulsifying agents, oils and waxes are melted and mixed thoroughly. The water-soluble ingredients dissolved in the water are warmed to a temperature slightly higher than the oil phase. The two phases then are mixed and stirred until cold. For convenience, but not necessity, the aqueous solution is added to the oil mixture. This method frequently is used in the preparation of ointments and creams.

Alternate Addition of the Two Phases to the Emulsifying Agent—A portion of the *oil*, if an O/W emulsion is being prepared, is added to all of the oil-soluble emulsifying agents with mixing, then an equal quantity of water containing all the water-soluble emulsifying agents is added with stirring until the emulsion is formed. Further portions of the oil and water are added alternately until the final product is formed. The high concentration of the emulsifying agent in the original emulsion makes the initial emulsification more likely and the high viscosity provides effective shearing action leading to small droplets in the emulsion. This method often is used successfully with soaps.

Multiple Emulsions—A recent innovation in emulsion technology is the development of multiple emulsions. The dispersed phase of these emulsions contains even smaller droplets which are miscible with the continuous phase. Thus, the multiple emulsion may be O/W/O where the aqueous phase is between two oil phases, or W/O/W where the internal and external aqueous phases are separated by an oil phase. In these systems both hydrophobic and hydrophilic emulsifiers are used and both have an effect on the yield and stability, as noted by Florence and Whitehill.[17]

It appears that O/W/O emulsions are formed better by lipophilic, nonionic surfactants using gum acacia-emulsified simple systems, while W/O/W multiple emulsions were formed better by nonionic surfactants in a two-stage emulsification procedure. A specific formulation for a W/O/W emulsion may be prepared by forming the primary (W/O) emulsion from isopropyl myristate (47.5%), sorbitan monooleate (2.5%) and distilled water (100%). This primary emulsion (50%) is added to a polyoxyethylene sorbitan monooleate (2% *w/v*) solution in water. While the technique of preparing these emulsions is more complicated, research indicates potential use of these emulsions for prolonged action, taste-masking, more effective dosage forms, parenteral preparations, protection against the external environment and enzyme entrapment.

Microemulsions—The coarse pharmaceutical macroemulsions appear white and tend to separate on standing. Microemulsions are translucent or transparent, do not separate and have a droplet diameter in the nanometer size range. The microemulsions are not always distinguishable from micellar solutions.

Both O/W and W/O types are possible and may be converted, one to the other, by adding more of the internal phase or by altering the type of emulsifier. As the internal phase is added, the emulsion will pass through a viscoelastic gel stage; with further addition, an emulsion of the opposite type will occur.

The most obvious benefit of microemulsions is their stability, thus providing dose uniformity. Usually, the emulsifier should be 20 to 30% of the weight of the oil used. The W/O systems are prepared by blending the oil and emulsifier with a little heat, if required, and then adding the water. The order of mixing for O/W systems is more flexible. One of the simplest methods is to blend the oil and the emulsifier and pour this into water with a little stirring. In no case can a microemulsion be formed unless there is a match between the oil and emulsifier.

If the emulsifier has been selected properly, microemulsification will occur almost spontaneously, leading to a satisfactory and stable preparation. The details of various preparations and the relationship between microemulsions and micellar solutions have been reviewed by Prince *et al.*[18] Microemulsions containing hydrocortisone have been prepared.

Equipment

When emulsions are prepared, energy must be expended to form an interface between the oily and aqueous phases. Emulsification equipment includes, therefore, a wide variety of agitators, homogenizers, colloid mills, jet mixers and ultrasonic devices. Griffin *et al.*,[16] Becher[19] and Peck *et al.*,[20] have evaluated the emulsification equipment used by pharmacists and drug manufacturers. These publications, along with journals such as *Pharmaceutical Technology*, should be consulted for further details on the use of such apparatus.

The preparation of emulsions on a large scale usually requires the expenditure of considerable amounts of energy for heating and mixing. Careful consideration of these processes has led to the development of low-energy emulsification by using an appropriate emulsification temperature and selective heating of the ingredients. This process, described by Lin,[21] involves the preparation of an emulsion concentrate subsequently diluted with the external phase at room temperature.

Agitators—Ordinary agitation or shaking may be used to prepare the emulsion. This method frequently is employed by the pharmacist, particularly in the emulsification of easily dispersed, low-viscosity oils. Under certain conditions, intermittent shaking is considerably more effective than ordinary continuous shaking. Continuous shaking tends to break up not only the phase to be dispersed but also the dispersion medium and, in this way, impairs the ease of emulsification. Laboratory shaking devices may be used for small-scale production.

The mortar and pestle are used widely by the prescription pharmacist in the extemporaneous preparation of emulsions. This equipment has very definite limitations because its usefulness depends largely on the viscidity of the emulsifying agent. A mortar and pestle cannot be used to prepare an emulsion if the emulsifying agent lacks viscidity (eg, gelatin solutions). These emulsifying agents will produce stable emulsions only if other types of equipment are used to mix the ingredients and the agent together.

Small electric mixers may be used to prepare emulsions at the prescription counter. They will save time and energy and produce satisfactory emulsions when the emulsifying agent is acacia or agar. However, the mixers cannot be used if the emulsifying agent is gelatin.

The commercially available *Waring Blendor* disperses efficiently by means of the shearing action of rapidly rotating blades. It transfers large amounts of energy and incorporates air into the emulsion. If an emulsion first is produced by using a blender of this type, the formulator must remember that the emulsion characteristics obtained in the laboratory will not be duplicated necessarily by the production-size agitators.

Production-size agitators include high-powered propeller-shaft stirrers immersed in a tank or self-contained units with

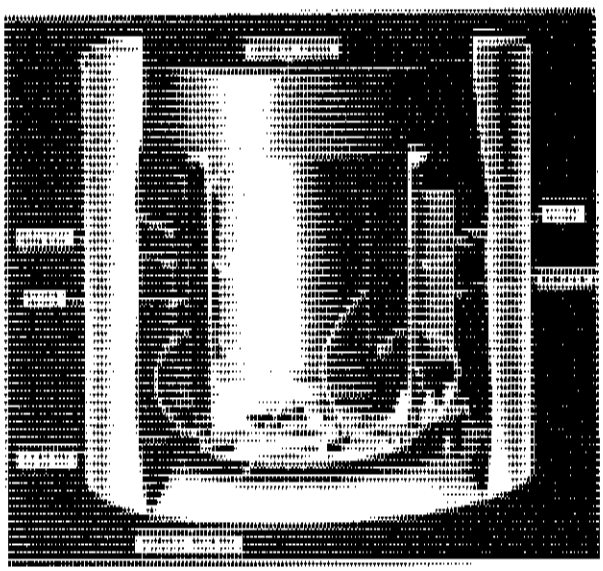
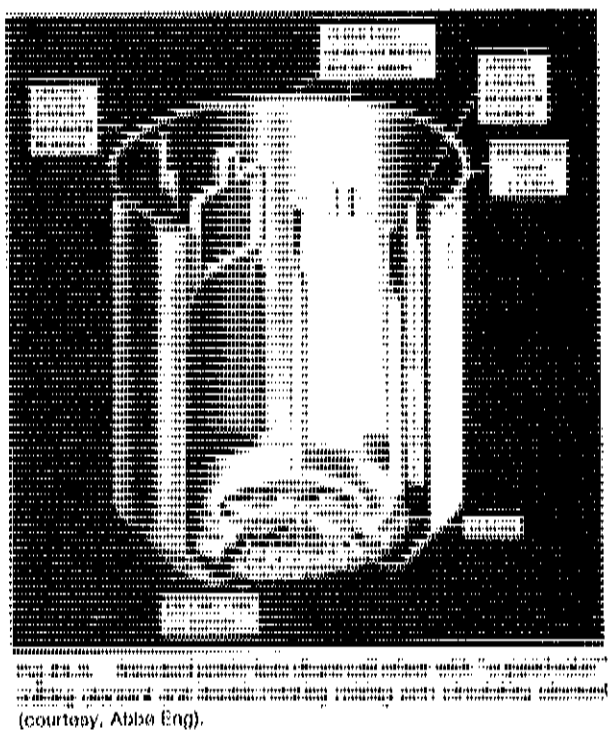


Fig 83-2. Standard slurry-type dispersall mixer with vaned-rotor "mixing" element and slotted draft-tube circulating element (courtesy, Abbe Eng).



(courtesy, Abbe Eng).

propeller and paddle systems. The latter usually are constructed so that the contents of the tank either may be heated or cooled during the production process. Baffles often are built into a tank and these increase the efficiency of agitation. Two mixers manufactured by the same company are shown in Figs 83-2 and 83-3.

Colloid Mills—The principle of operation of the colloid mill is the passage of the mixed phases of an emulsion formula between a stator and a high-speed rotor revolving at speeds of 2000 to 18,000 rpm. The clearance between the rotor and the stator is adjustable, usually from 0.001 in upward. The emulsion mixture, in passing between the rotor and stator, is subjected to a tremendous shearing action which effects a fine dispersion. A colloid mill and various rotors are shown in Figs 83-4 and 83-5. The operating principle is the same for all, but each manufacturer incorporates specific features which result in changes in operating efficiency. The shearing forces applied in the colloid mill may result in a temperature increase within the emulsion. It may be necessary, therefore, to cool the equipment when the emulsion is being produced.

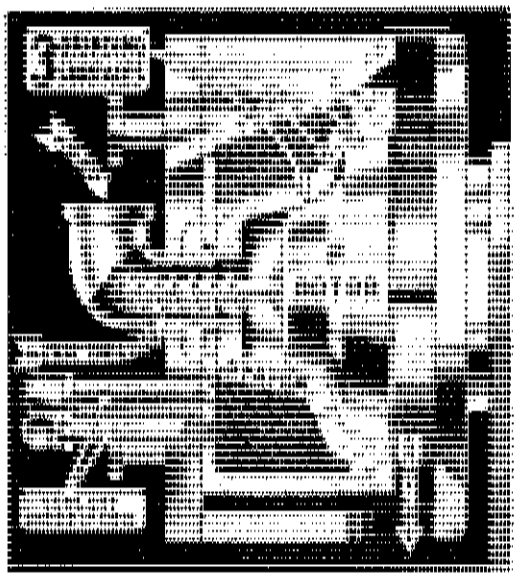


Fig 83-4. A colloid mill shown in cross section (courtesy, Tri-Homo).

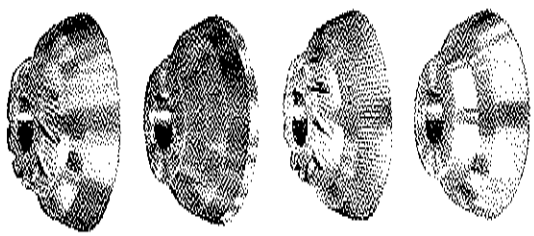


Fig 83-5. Types of rotors used in colloid mills. These may be smooth (for most emulsions), serrated (for ointments and very viscous products) or of vitrified stone (for the paints and pigment dispersions) (courtesy, Tri-Homo).

Homogenizers and Viscolizers—In these two types of equipment the mixed phases are passed between a finely ground valve and seat under high pressure. This, in effect, produces an atomization which is enhanced by the impact received by the atomized mixture as it strikes the valve head. They operate at pressures of 1000 to 5000 psi and produce some of the finest dispersions obtainable in an emulsion.

Homogenizers may be used in one of two ways:

1. The ingredients in the emulsion are mixed and then passed through the homogenizer to produce the final product.
2. An emulsion is prepared in some other way and then passed through a homogenizer for the purpose of decreasing the particle size and obtaining a greater degree of uniformity and stability.

Two-stage homogenizers are constructed so that the emulsion, after treatment in the first valve system, is conducted directly to another where it receives a second treatment. A single homogenization may produce an emulsion which, although its particle size is small, has a tendency to clump or form clusters. Emulsions of this type exhibit increased creaming tendencies. This is corrected by passing the emulsion through the first stage of homogenization at a high

pressure (eg, 3000 to 5000 psi) and then through the second stage at a greatly reduced pressure (eg, 1000 psi). This breaks down any clusters formed in the first step.

For small-scale extemporaneous preparation of emulsions, the inexpensive *hand homogenizer* (available from *Med Times*) is particularly useful. It is probably the most efficient emulsifying apparatus available to the prescription pharmacist. The two phases, previously mixed in a bottle, are hand pumped through the apparatus. Recirculation of the emulsion through the apparatus will improve its quality.

A homogenizer does not incorporate air into the final product. Air may ruin an emulsion because the emulsifying agent is adsorbed preferentially at the air/water interface, followed by an irreversible precipitation termed *denaturization*. This is particularly prone to occur with protein emulsifying agents.

Homogenization may spoil an emulsion if the concentration of the emulsifying agent in the formulation is less than that required to take care of the increase in surface area produced by the process.

The temperature rise during homogenization is not very large. However, temperature does play an important role in the emulsification process. An increase in temperature will reduce the viscosity and, in certain instances, the interfacial tension between the oil and the water. There are, however, many instances, particularly in the manufacturing of cosmetic creams and ointments, where the ingredients will fail to emulsify properly if they are processed at too high a temperature. Emulsions of this type are processed first at an elevated temperature and then homogenized at a temperature not exceeding 40°.

Figure 83-6 shows the flow through the homogenizing valve, the heart of the high-pressure APV Gaulin homogenizer. The product enters the valve seat at high pressure, flows through the region between the valve and the seat at high velocity with a rapid pressure drop and then is discharged as a homogenized product. It is postulated that circulation and turbulence are responsible mainly for the homogenization that takes place. Different valve assemblies, two stage valve assemblies and equipment with a wide range of capacities are available.

The Macro Flow-Master *Kom-bi-nator* employs a number of different actions, each of which takes the ingredients a little further along in the process of subdividing droplets, until complete homogenization results. The machine is equipped with a pump which carries the liquid through the various stages of the process. In the first stage, the ingredients are forced between two specially designed rotors (gears) which shoot the liquid in opposite directions in a small chamber and, in this way, are mixed thoroughly. These rotors also set up a swirling action in the next chamber into

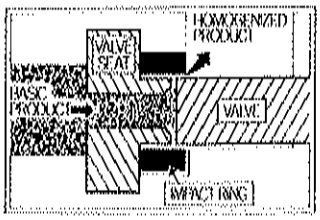


Fig 83-6. Operation of the homogenizer valve assembly (Courtesy APV Gaulin).

which the liquid is forced and swirled back and forth in eddies and crosscurrents. The second stage is a pulsing or vibrating action at rapid frequency. The product then leaves this chamber, goes through a small valve opening and is dashed against the wall of the homogenizing chamber. Pressure is applied, but it is not as great as that used in other types of homogenizers. Pressure is controlled accurately by adjusting devices on the front of the machine, and temperature is controlled by passing coolants through the stators.

Ultrasonic Devices—The preparation of emulsions by the use of ultrasonic vibrations also is possible. An oscillator of high frequency (100 to 500 kHz) is connected to two electrodes between which is placed a piezoelectric quartz plate. The quartz plate and electrodes are immersed in an oil bath and, when the oscillator is operating, high-frequency waves flow through the fluid. Emulsification is accomplished by simply immersing a tube containing the emulsion ingredients into this oil bath. Considerable research has been done on ultrasonic emulsification, particularly with regard to the mechanism of emulsion formation by this method. Limited data indicate that these devices will produce stable emulsions only with liquids of low viscosity. The method is not practical, however, for large-scale production of emulsions.

Special techniques and equipment in certain instances, will produce superior emulsions, including rapid cooling, reduction in particle size or ultrasonic devices. A wide selection of equipment for processing both emulsions and suspensions has been described by Eisberg.[22] A number of improvements have been made to make the various processes more effective and energy-efficient.

General methods are available for testing the instability of emulsions including bulk changes, centrifugal and ultracentrifugal studies, dielectric measurement, surface-area measurement and accelerated-motion studies. Low-shear rheological studies measuring viscoelasticity are suggested as the optimal method of stability testing.

Suspensions

The physical chemist defines the word "suspension" as a two-phase system consisting of a finely divided solid dispersed in a solid, liquid or gas. The pharmacist accepts this definition and can show that a variety of dosage forms fall within the scope of the preceding statement. There is, however, a reluctance to be all-inclusive, and it is for this reason that the main emphasis is placed on solids dispersed in liquids. In addition, and because there is a need for more specific terminology, the pharmaceutical scientist differentiates between such preparations as suspensions, mixtures, magmas, gels and lotions. In a general sense, each of these preparations represents a suspension, but the state of subdivision of the insoluble solid varies from particles which settle gradually on standing to particles which are colloidal in nature. The lower limit of particle size is approximately 0.1 μm, and it is the preparations containing dispersed solids of this magnitude or greater that are defined pharmaceutically as suspensions.

Certain authors also include liniments, and the newer sustained-release suspensions, in any discussion of this particular subject. The former preparations now usually are considered as solutions although a number of older liniments were, in fact, suspensions. The sustained-release suspensions represent a very specialized class of preparation and, as such, are discussed in more detail in Chapter 91. Some insoluble drugs also are administered in aerosol form; one example is dexamethasone phosphate suspended in a propellant mixture of fluorochlorocarbons. More detail on aerosols is available in Chapter 92.

Suspension formulation and control is based on the prin-

ciples outlined in Chapters 19 and 20. Formulation involves more than suspending a solid in a liquid. A knowledge of the behavior of particles in liquids, of suspending agents and of flavors and colors is required to produce a satisfactory suspension.

Briefly, the preparation of a stable suspension depends upon the appropriate dispersion of the drug in the suspending medium. To ensure that the particles are wetted by the dispersion medium, a surface-active agent should be used, especially if the dispersed phase is hydrophobic. The suspending agent in the aqueous medium then can be added. Alternatively, the dry suspending agent can be mixed thoroughly with the drug particles and then triturated with the diluent. Other approaches to suspension preparation include the formation of a flocculated suspension and also a flocculated preparation in a suspending vehicle. Details of these procedures are given in Chapter 19.

The most efficient method of producing fine particles is by dry milling prior to suspension. Suspension equipment such as colloid mills or homogenizers normally are used in wet-milling finished suspensions to reduce particle agglomerates. These machines (Fig 83-4) usually have a stator and a rotor which effects the dispersion action. Several methods of producing small uniform dry particles are micropulverization fluid-energy grinding, spray-drying and controlled precipitation with ultrasound as described by Nash.[33]

The choice of an appropriate suspending agent depends upon the use of the products (external or internal), facilities for preparation and the duration of storage.

Preparations made extemporaneously for internal use may include, as suspending agents, acacia, methylcellulose or other cellulose derivatives, sodium alginate or tragacanth.

Extemporaneous preparations of suspensions for internal use showing good flow and suspending properties are provided by sodium carboxymethylcellulose 2.5%, tragacanth 1.25% and guar gum 0.5%. Avicel RC-591, a coprecipitate of microcrystalline cellulose and sodium carboxymethylcellulose stabilized with hydroxypropyl methylcellulose, has been used as a suspending vehicle for propranolol and orphenadrine hydrochloride dispersions prepared from tablets. It also may serve as a general-purpose suspending agent. Carbopol 934, 0.3% or greater, was a satisfactory suspending agent for sulfamethazine 10%, maintaining a permanent suspension for more than 6 months.

Agents suitable for external use include bentonite, methylcellulose or other cellulose derivatives, sodium alginate or tragacanth. Agents which may require high-speed equipment and which are suitable for internal or external use include aluminum magnesium silicates and carbomer.[34]

Preparations such as those mentioned above possess certain advantages over other dosage forms. Some drugs are insoluble in all acceptable media and, therefore, must be administered as a solid, nonsolution dosage form (tablet, capsule, etc), or as a suspension. Because of its liquid character, the last preparation insures some uniformity of dosage but does present some problems in maintaining a consistent dosage regimen. Disagreeable tastes can be covered by using a suspension of the drug or a derivative of the drug, an example of the latter being chloramphenicol palmitate. Suspensions prepared from ion-exchange resins containing an ionic drug can be used not only to minimize the taste of the drug but also to produce a prolonged-action product, since the drug is exchanged slowly for other ions within the gastrointestinal tract.

Suspensions also are chemically more stable than solutions. This particularly is important with certain antibiotics, and the pharmacist often is called on to prepare such a suspension just prior to dispensing the medication. In addition, a suspension is an ideal dosage form for patients who have difficulty swallowing tablets or capsules, which is particularly important in administering drugs to children. An alternate method to enhance compliance includes flavored nystatin "popsicles" which can be prepared by freezing a suspension of the drug so that the taste is improved during the treatment of oral candidiasis.

Suspensions should possess certain basic properties. The dispersed phase should settle slowly and be redispersed readily on shaking. They should not cake on settling and the viscosity should be such that the preparation pours easily. As with all dosage forms, there should be no question as to the chemical stability of the suspension. Appropriate preservatives should be incorporated in order to minimize microbiological contamination. The suspension must be acceptable to the patient on the basis of its taste, color and cosmetic qualities (elegance), the latter two factors being of particular importance in preparations intended for external use.

Gels

Pharmaceutical terminology is, at best, confusing and no two authors will classify gels, jellies, magmas, milks and mixtures in the same way. The NF described Gels as a special class of pharmaceutical preparations but considered Jellies under the same heading. The latter preparations usually contain water-soluble active ingredients and, therefore, are considered in another part of this chapter. The USP definition for Gels is

> Gels are semisolid systems of either suspensions made up of small inorganic particles or large organic molecules interpenetrated by a liquid. Where the gel mass consists of a network of small discrete particles, the gel is classified as a two-phase system (eg, Aluminum Hydroxide Gel). In a two-phase system, if the particle size of the dispersed phase is relatively large, the gel mass sometimes is referred to as a magma (eg Bentonite Magma). Both gels and magmas may be thixotropic, forming semisolids on standing and becoming liquid on agitation. They should be shaken before use to ensure homogeneity and should be labeled to that effect.
>
> Single-phase gels consist of organic macromolecules distributed uniformly throughout a liquid in such a manner that no apparent boundaries exist between the dispersed macromolecules and the liquid. Single-phase may be made from synthetic macromolecules (eg, Carbomer) or from natural gums (eg, Tragacanth). The latter preparations also are called mucilages. Although these gels are commonly aqueous, alcohol and oils may be used as the continuous phase. For example, mineral oil can be combined with a polyethylene resin to form an oleaginous ointment base.

The USP states that each 100 g of Aluminum Hydroxide Gel contains the equivalent of not less than 3.6 and not more than 4.4 g of aluminum oxide (Al_2O_3), in the form of aluminum hydroxide and hydrated oxide, and it may contain varying quantities of basic aluminum carbonate and bicarbonate. The gel itself usually is prepared by the interaction of a soluble aluminum salt, such as a chloride or sulfate, with ammonia solution, sodium carbonate or bicarbonate. The reactions which occur during the preparation are

$$3CO_3^{2-} + 3H_2O \rightarrow 3HCO_3^- + 3OH^-$$

$$[Al(H_2O)_6]^{3+} + 3OH^- \rightarrow [Al(H_2O)_3(OH)_3] + 3H_2O$$

$$2HCO_3^- \rightarrow CO_3^{2-} + H_2O + CO_2$$

The physical and chemical properties of the gel will be affected by the order of addition of reactants, pH of precipitation, temperature of precipitation, concentration of the reactants, the reactants used and the conditions of aging of the precipitated gel.

Aluminum Hydroxide Gel is soluble in acidic (or very strongly basic) media. The mechanism in acidic media is

$$\text{Aluminum Hydoxide Gel} + 3H_3O^+ \rightarrow [Al(H_2O)_3(OH)_3]^0$$

$$[Al(H_2O)_3]^0 + H_3O^+ \rightarrow [Al(H_2O)_4(OH)_2]^+ + H_2O$$

$$[Al(H_2O)_4(OH)_2]^{+} + H_3O^{+} \rightarrow [Al(H_2O)_5(OH)]^{2+} + H_2O$$

$$[Al(H_2O)_5(OH)]^{2+} + H_3O^{+} \rightarrow [Al(H_2O)_6]^{3+} + H_2O$$

It is unlikely that the last reaction given proceeds to completion. Since the activity of the gel is controlled by its insolubility (solubility will decrease with an increase in the pH of the gastric media), there is no acid rebound. Further, since a certain quantity of insoluble gel always is available, the neutralizing capability of the gel extends over a considerable period of time.

Aluminum hydroxide gels also may contain peppermint oil, glycerin, sorbitol, sucrose, saccharin and various preservatives. Sorbitol improves the acid-consuming capacity, apparently by inhibiting a secondary polymerization that takes place on aging. In addition, polyols such as mannitol, sorbitol and inositol have been shown to improve the stability of aluminum hydroxide and aluminum hydroxycarbonate gels.

Aluminum Hydroxide and Belladonna Mixture PC

Belladonna Tincture	100 mL
Chloroform Spirit	50 mL
Aluminum Hydroxide Gel to	1000 mL

It should be noted, however, that the addition of other drugs (eg, antibiotics) to the gel may result in a loss of the activity anticipated for that active ingredient.

Generally, if left undisturbed for some time, gels may become semisolid or gelatinous. With some gels, small amounts of water may separate on standing.

The single-phase gels are being used more frequently in pharmacy and cosmetics because of several properties: semisolid state, high degree of clarity, ease of application and ease of removal and use. The gels often provide a faster release of drug substance, independent of the water solubility of the drug, as compared to creams and ointments. Some drugs used in medication gels include urea, hydrogen peroxide, ephedrine sulphate, erythromycin and povidone-iodine.

Gels may be used as lubricants for catheters, bases for patch testing, sodium chloride gels for electrocardiography, fluoride gels for topical dental use and for intravaginal administration (prostaglandin-E_2 gel).

Gels can be prepared from a number of pharmaceutical agents such as tragacanth 2 to 5%, sodium alginate 2 to 10%, gelatin 2 to 15%, methylcellulose 2 to 4%, sodium carboxymethyl-cellulose 2 to 5%, carbomer 0.3 to 5% or polyvinyl alcohols 10 to 20% as noted by Carter.[25] Other gelling agents include methylhydroxyethyl cellulose, polyoxyethylene-polyoxypropylene, hydroxyethyl cellulose and gelatin. Gels prepared from nonpolar materials such as magnesium soap-hydrocarbon and hydrocarbons are being investigated.

The percentages above indicate the concentration ranges of the gelling agent. The lower-percentage preparations may be used as lubricants and the higher-percentage preparations as dermatological bases. Some of the gelling agents are available in different grades indicating the viscosity at a definite concentration. In general, high-viscosity grades result in gels at lower concentrations.

Gels recently have been prepared in adhesive form in order to increase the contact time of the active ingredients, such as insulin with the oral and nasal mucosa, leading to a decrease in plasma glucose. This system also has been investigated as a vaginal dosage form for cervical cancer and a topical dosage form for aphthous stomatitis.

Preservatives should be incorporated into the gels, especially those prepared from natural sources. Appropriate preservatives, depending upon use and the gelling agent, include the parabens at about 0.2%, benzoic acid 0.2% (if the product is acidic) and chlorocresol 0.1%.

The preparation of a few gel bases is given below:

Sodium Alginate Gel Base

Sodium Alginate	2–10 g
Glycerin	2–10 g
Methyl Hydroxybenzoate	0.2 g
a soluble calcium salt (calcium or gluconate)	0.5 g
Purified Water, to make	100 mL

The sodium alginate is wetted in a mortar with glycerin, which aids the dispersion. The preservative is dissolved in about 80 mL of water with the aid of heat, allowed to cool and the calcium salt added, which will increase the viscosity of the preparation. This solution is stirred in a high speed stirrer and the sodium alginate-glycerin mixture added slowly while stirring, until the preparation is homogeneous. The preparation should be stored in a tightly sealed container in a wide mouth jar or tube.

Carbomer Jelly

Carbopol 934	2 g
Triethanolamine	1.05 mL
Parabens	0.2 g
Purified Water, to make	100 mL

The parabens are dissolved in 95 mL of water with the aid of heat and allowed to cool. The Carbopol 934, a commercial grade of carbomer, is added in small amounts to the solution using a high speed stirrer and, after a smooth dispersion is obtained, the preparation is allowed to stand permitting entrapped air to separate. Then the gelling agent, triethanolamine, is added, dropwise, stirring with a plastic spatula to avoid entrapping air and the remaining water incorporated. Other concentrations of carbomer can be used to prepare gels, creams or suspensions.

The USP lists a number of gels: Sodium Fluoride and Phosphoric Acid Gel for application to the teeth to reduce cavities, Betamethasone Benzoate Gel and Fluocinonide Gel, anti-inflammatory corticosteroids, Tolnaftate Gel, an antifungal agent and Tretinoin Gel for the treatment of acne. Refer to the specific monographs in this text for more information.

Lotions

Lotions usually are liquid suspensions or dispersions intended for external application to the body. They may be prepared by triturating the ingredients to a smooth paste and then adding the remaining liquid phase cautiously. High-speed mixers or homogenizers produce better dispersions and, therefore, are the tools of choice in the preparation of larger quantities of lotion. Calamine Lotion USP is the classic example of this type of preparation and consists of finely powdered, insoluble solids held in more or less permanent suspension by the presence of suspending agents and/or surface-active agents. Many investigators have studied Calamine Lotion and this has led to the publication of many formulations, each possessing certain advantages over the others but none satisfying the collective needs of all dermatologists.

Phenolated Calamine Lotion USP contains 10 mL of liquefied phenol in sufficient calamine lotion to make the product measure 1000 mL. Formulations containing Avicel R (hydrated microcrystalline cellulose, *FMC*) and carboxymethylcellulose settle less than the official preparations.

Calamine Lotion

Calamine	8 g
Zinc Oxide	8 g
Glycerin	2 mL
Avicel R Gel	2 g
Carboxymethylcellulose	2 g
Calcium Hydroxide Solution, a sufficient quantity, to make	100 mL

Mix 45 g of Avicel R with 55 g of water with a suitable electric mixer. This gel is used in the preparation of the calamine lotion. Mix the calamine and the zinc oxide with the glycerin, the gel and the carboxymethylcellulose. Add sufficient calcium hydroxide solution to make the product measure 100 mL.

To prepare Phenolated Calamine Lotion add 1 mL of Liquefied Phenol during the mixing stage.

Suspensions also may be formed by chemical interaction in the liquid. White Lotion is an example.

White Lotion

Zinc Sulfate	40 g
Sulfurated Potash	40 g
Purified Water, a sufficient quantity (to make	1000 mL

Dissolve the zinc sulfate and the sulfurated potash separately, each in 450 mL of purified water and filter each solution. Add slowly the sulfurated potash solution to the zinc sulfate solution with constant stirring. Then add the required amount of purified water, and mix.

Sulfurated potash is a solid of variable composition but usually is described as $K_2S_5 \cdot K_2S_2O_3$. The chemical reaction which occurs when sulfurated potash solution is added to the zinc sulfate is

$$ZnSO_4 \cdot 7H_2O + K_2S_5 \cdot K_2S_2O_3 \rightarrow ZnS + S_2 + K_2SO_4 + K_2S_2O_3 + 7H_2O$$

This lotion must be prepared fresh and does not contain a suspending agent. Bentonite Magma has been used in some formulations. Coffman and Huyck[26] include a detailed discussion of the chemistry and the problems involved in the preparation of a suitable product.

The USP recognizes a second type of lotion. These are emulsions of the O/W type stabilized by a surface-active agent. Benzyl Benzoate Lotion is an example. Some lotions are clear solutions and, in fact, the active ingredient of one official lotion, Dimethisoquin Hydrochloride Lotion USP XX is a water-soluble substance. However, one unofficial formulation for this lotion lists dimethisoquin hydrochloride, menthol and zinc oxide as active ingredients and the preparation thus becomes a suspension. Several lotions are listed in the USP and contain, for example, antibiotics, steroids, keratolytics and scabicides.

A formula for hydrocortisone lotion is given in the PC.

Hydrocortisone Lotion

Hydrocortisone, in ultrafine powder	10.0 g
Chlorocresol	0.5 g
Self-emulsifying monostearin	40.0 g
Glycerol	63.0 g
Purified water, freshly boiled and cooled to make	1000.0 g

To prepare the base, the chlorocresol is dissolved in 850 mL of water with the aid of gentle heat, the self-emulsifying monostearin is added and the mixture heated to 60° with stirring until completely dispersed. The hydrocortisone is triturated with the glycerol and the trituration is then incorporated, with stirring, into the warm base, allowed to cool while stirring, then added the remainder of the water and mixed.

Lotions usually are applied without friction. Even so, the insoluble matter should be divided very finely. Particles approaching colloidal dimensions are more soothing to inflamed areas and effective in contact with infected surfaces. A wide variety of ingredients may be added to the preparation to produce better dispersions or to accentuate its cooling, soothing, drying or protective properties. Bentonite is a good example of a suspending agent used in the preparation of lotions. Methylcellulose or sodium carboxymethylcellulose will localize and hold the active ingredient in contact with the affected site. A formulation containing glycerin will keep the skin moist for a considerable period of time. The drying and cooling effect may be accentuated by adding alcohol to the formula.

Dermatologists frequently prescribe lotions containing anesthetics, antipruritics, antiseptics, astringents, germicides, protectives or screening agents, to be used in treating or preventing various types of skin diseases and dermatitis. Antihistamines, benzocaine, calamine, resorcin, steroids, sulfur, zinc oxide, betamethasone derivatives, salicylic acid, safflower oil, minoxidil and zirconium oxide are ingredients common in unofficial lotions. In many instances the cosmetic aspects of the lotion are of great importance. Many lotions compare badly with cosmetic preparations of a similar nature. The manufacture of fine lotions to meet the specialized needs of the dermatologist provides the pharmacist with an excellent opportunity to demonstrate his professional competence. Recent extensive studies on lotions, as described by Harb,[27] will assist the pharmacist to attain this goal.

Lotions tend to separate or stratify on long standing, and they require a label directing that they be shaken well before each use. All lotions should be labeled "For External Use Only."

Microorganisms may grow in certain lotions if no preservative is included. Care should be taken to avoid contaminating the lotion during preparation, even if a preservative is present.

Magmas and Milks

Magmas and milks are aqueous suspensions of insoluble, inorganic drugs and differ from gels mainly in that the suspended particles are larger. When prepared, they are thick and viscous and, because of this, there is no need to add a suspending agent.

Bentonite Magma USP is prepared by simple hydration. Two procedures are given in the compendium for the preparation of this product.

Magmas also may be prepared by chemical reaction. Magnesium hydroxide is prepared by the hydration of magnesium oxide.

$$MgO + H_2O \rightarrow Mg(OH)_2$$

Milk of Magnesia USP is a suspension of magnesium hydroxide containing 7.0–8.5% $Mg(OH)_2$. It has an unpleasant, alkaline taste which can be masked with 0.1% citric acid (to reduce alkalinity) and 0.05% of a volatile oil or a blend of volatile oils.

Milk of Bismuth contains bismuth hydroxide and bismuth subcarbonate in suspension in water. The Magma is prepared by reacting bismuth subnitrate with nitric acid and ammonium carbonate with ammonia solution and then mixing the resulting two solutions.

The following reactions occur during the preparation of the magma.

$$(NH_4)_2CO_3 \rightarrow 2NH_4^+ + CO_3^{2-}$$

$$NH_3 + H_2O \rightarrow NH_4^+ + OH^-$$

$$2BiO^+ + CO_3^{2-} \rightarrow (BiO)_2CO_3$$

$$BiO^+ + OH^- \rightarrow BiO(OH)$$

If the insoluble substance is precipitated fresh by mixing hot, dilute solutions, there is only slight sedimentation on standing. This characteristic of magmas sometimes is enhanced by passing the product through a colloid mill.

For the most part, magmas and milks are intended for internal use, although Bentonite Magma is used primarily as a suspending agent for insoluble substances eg, Milk of Magnesia USP and Dihydroxy Aluminum Aminoacetate Magma USP, either for local application or for internal use. All magmas require a "Shake Well" label. Freezing must be avoided.

Several antimicrobial preservatives have been tested in liquid antacid preparations for their stability and effectiveness, such as benzoic acid, chlorhexidine, methylparaben,

propylparaben, sorbic acid, propylene glycol or ethanol. It was found that a combination of methylparaben and sorbic acid was superior to the parabens alone.

Mixtures

The official mixtures are aqueous, liquid preparations which contain suspended, insoluble, solid substances and are intended for internal use. The insoluble substance does not make the mixture very viscous, and the particles may be held in suspension by using suitable suspending or thickening agents. This class was introduced originally to secure uniformity in the formulas of certain well-known and largely used preparations. Frequently, the term *mixture* is applied loosely to aqueous preparations of every description. The term *shake mixture* is used often for liquid preparations which contain insoluble ingredients and, therefore, must be shaken before use. The USP does not recognize the term. The term *suspension* now is used to describe a number of similar preparations. The PC uses the term *mixtures* and includes suspensions in this category, for example:

Ammonium Chloride Mixture PC

Ammonium Chloride	100 g
Aromatic Ammonia Solution	50 mL
Liquorice Liquid Extract	100 mL
Water, for preparations to	1000 mL

It should be prepared recently.

The term mixture occurs in the expression dry mixture, which may be used to describe many USP products, in particular, antibiotic powders for oral solutions which are described on page 1527.

The pectin and the tragacanth in Kaolin Mixture with Pectin (page 796) act as suspending agents. An alternate formula, based on Veegum (*Vanderbilt*) and sodium carboxymethylcellulose, has been proposed by Kalish.[28]

Kaolin Mixture with Pectin

Veegum	0.88 g
Sodium Carboxymethylcellulose	0.22 g
Purified Water	79.12 g
Kaolin	17.50 g
Pectin	0.44 g
Saccharin	0.00 g
Glycerin	1.75 g

Add the Veegum and the sodium carboxymethylcellulose to the water with continuous stirring. Add, with mixing, the kaolin. Mix the pectin, saccharin and glycerin and add to the suspension. A preservative and flavoring agent may be added to the product.

The insoluble material in mixtures must be in a very finely divided state and uniformly distributed throughout the preparation. This is accomplished with colloid mills, special methods of precipitation and suspending agents. There are three main reasons for having the insoluble substances in as fine a state of subdivision as possible.

1. The more nearly the colloidal state is approached by protectives, such as kaolin, magnesium trisilicate or magnesium phosphate, the more active they become as adsorbents and protectives when in contact with inflamed surfaces.
2. Finely divided particles are suspended more readily and settle out much more slowly than large particles, thus enabling the patient to obtain uniform doses of suspended substances. Homogeneous mixtures are desirable especially when administering medication to form an evenly distributed, protective coating on the gastrointestinal tract.
3. The palatability of many preparations is enhanced by the use of colloidal suspending agents.

Mixtures containing suspended material should have a "Shake Well" label affixed to the container in which they are dispensed.

Mixtures, including suspensions, are subject to contamination by microorganisms that remain viable and are a potential health hazard during the period of use of the products. Survival times of organisms depend on the preservative used. A kaolin pediatric mixture that contains benzoic acid kills organisms rapidly, whereas organisms survived for more than a week in a magnesium trisilicate mixture that contained no more than a trace of peppermint oil, as noted by Westwood.[29]

Occasionally, it is necessary to prepare suspensions from crushed tablets. A general formula for this purpose is given.[24]

Methylcellulose 20	0.75
Parabens	0.1
Purified Water	60.0
Propylene Glycol	2.0
Simple Syrup, to make	100.0

An extemporaneous suspension of cimetidine tablets which retained its potency at 40° over 14 days is:

Cimetidine 300-mg tablets	24 (7.2 g)
Glycerin	10 mL
Simple Syrup, to make	120 mL

The tablets are triturated to a fine powder using a mortar, the mixture is levigated with the glycerin, simple syrup added, mixed well, placed in a blender until smooth and then refrigerated.[30]

Satisfactory suspensions have been compounded from diazepam tablets and propranolol hydrochloride tablets, and they possess chemical stability for 60 days and 4 months, respectively, at room temperature or under refrigeration. Frequently, since the drug may be soluble, it is the excipients which are being suspended.

A comprehensive checklist of suspension formulations has been reported in the literature by Scheer.[31]

Official Suspensions

The USP places particular emphasis on the term suspension by providing specific definitions for a variety of oral, parenteral and ophthalmic preparations formulated in such a way that an insoluble substance is suspended in a liquid at some stage of the manufacturing or dispensing process. The USP definition begins as follows:

Suspensions are preparations of finely divided, undissolved drugs dispersed in liquid vehicles. Powders for suspension are preparations of finely powdered drugs intended for suspension in liquid vehicles. An example of the ready-to-use type is *Trisulfapyrimidines Oral Suspension*, in which the three sulfapyrimidines are already suspended in a liquid flavored vehicle in a form suitable for oral administration. *Tetracycline for Oral Suspension* is finely divided tetracycline mixed with suspending and dispersing agents. It is intended to be constituted with the prescribed volume of purified water and mixed before it is dispensed by the pharmacist for oral administration to the patient.

Neither this definition nor the monographs give specific directions for the preparation of the suspension, although pharmacopeias usually permit the addition of suitable flavoring agents, suspending agents, preservatives and certified color additives. One procedure for the preparation of the commonly used Trisulfapyrimidines Oral Suspension is given below.

Trisulfapyrimidines Oral Suspension

Veegum	1.00 g
Syrup USP	90.60 g
Sodium Citrate	0.78 g
Sulfadiazine	2.54 g
Sulfamerazine	2.54 g
Sulfamethazine	2.54 g

Add the Veegum, slowly and with continuous stirring, to the syrup. Incorporate the sodium citrate into the Veegum-syrup mixture. Premix the salts drugs, add to the syrup, stir and homogenize. Add sufficient 5% citric acid to adjust the pH of the product to 5.6. A preservative and a flavoring agent may be added to the product.

Methods of preparation for those formulations which contain several active ingredients and are produced in large quantities tend to be more complex than that given above.

Many formulations for suspensions are given in the PC under *Mixtures*.

A properly prepared suspension has a number of desirable properties:

1. The suspended material should not settle rapidly.
2. Particles that do settle should not form a hard cake and easily should be resuspended uniformly on shaking.
3. The suspension should pour freely from the container.

Insoluble powders that do not disperse evenly throughout the suspending medium, when shaken, should be powdered finely and levigated with a small amount of an agent such as glycerin, alcohol or a portion of the dispersion of the suspending agent. The other ingredients are incorporated and the remainder of the dispersion of the suspending agent is incorporated gradually by trituration to produce the appropriate volume.

Suspensions intended for parenteral or ophthalmic use also are described in the USP. For a discussion of these suspensions, see Chapter 84 and 86.

Extracts

Extraction

Extraction, as the term is used pharmaceutically, involves the separation of medicinally active portions of plant or animal tissues from the inactive or inert components by using selective solvents in standard extraction procedures.

The products so obtained from plants are relatively impure liquids, semisolids or powders intended only for oral or external use. These include classes of preparations known as decoctions, infusions, fluidextracts, tinctures, pilular (semisolid) extracts and powdered extracts. Such preparations popularly have been called galenicals, after Galen, the 2nd century Greek physician. For additional information concerning extraction and extractives, see RPS 15, Chapter 86.

Extraction continues to be of considerable interest in order to obtain improved yields of drugs derived from plant and animal sources. For example, improved extraction of digitalis glycosides has been carried out using a pulsating, perforated, bottom column. Other techniques include ultrasonics, rotary-film evaporators, liquid and supercritical carbon dioxide, hydrodistillation, liquid chromatography, multiple-solvent extraction, countercurrent extraction and gravitation dynamics.

In this discussion we are concerned primarily with basic extraction procedures for crude drugs to obtain the therapeutically desirable portion and eliminate the inert material by treatment with a selective solvent, known as the menstruum. Extraction differs from solution in that the presence of insoluble matter is implied in the former process. The principal methods of extraction are maceration, percolation, digestion, infusion and decoction. The quality of the finished product can be enhanced by standardizing primary extracts and carrying out analytical assays during production on the raw materials, intermediate products and manufacturing procedures.

The processes of particular importance, insofar as the USP is concerned, are those of maceration and percolation. Most pharmacopeias refer to such processes for extraction of active principles from crude drugs.

Maceration—In this process the solid ingredients are placed in a stoppered container with the whole of the solvent and allowed to stand for a period of at least 3 days, with frequent agitation, until soluble matter is dissolved. The mixture then is strained, the marc (the damp solid material) pressed and the combined liquids clarified by filtration or by decantation, after standing.

Percolation—This is the procedure used most frequently to extract the active ingredients in the preparation of tinctures and fluidextracts. Certain specific procedural details are provided in the USP, which should be consulted for such information. In the PC general procedure, a percolator (a narrow, cone-shaped vessel open at both ends) is used. The solid ingredient(s) are moistened with an appropriate amount of the specified menstruum and allowed to stand for approximately 4 hr in a well-closed container, after which the drug mass is packed into the percolator. Sufficient menstruum is added to saturate the mass and the top of the percolator is closed. When the liquid is about to dip from the neck (bottom) of the percolator, the outlet is closed. Additional menstruum is added to give a shallow layer above the mass, and the mixture is allowed to macerate in the closed percolator for 24 hr. The outlet of the percolator then is opened and the liquid contained therein allowed to drip slowly, additional menstruum being added as required, until the percolate measures about three-quarters of the required volume of the finished product. The marc is pressed and the expressed liquid is added to the percolate. Sufficient menstruum is added to produce the required volume, and the mixed liquid clarified by filtration or by allowing it to stand and then decanting.

Digestion—This is a form of maceration in which *gentle heat* is used during the process of extraction. It is used when moderately elevated temperature is not objectionable and the solvent efficiency of the menstruum is increased thereby.

Infusion—An infusion is a dilute solution of the readily soluble constituents of crude drugs. Fresh infusions are prepared by macerating the drugs for a short period of time with either cold or boiling water. US official compendia have not included infusions for some time. An example is Concentrated Compound Gentian Infusion BP 1973.

Decoction—This once-popular process extracts water-soluble and heat-stable constituents from crude drugs by boiling in water for 15 min, cooling, straining and passing sufficient cold water through the drug to produce the required volume.

Extractive Preparations

After a solution of the active constituents of a crude drug is obtained by maceration or percolation, it may be ready for use as a medicinal agent, as with certain tinctures or fluidextracts, or it may be processed further to produce a solid or semisolid extract.

For a discussion of *resins* and *oleoresins* obtained by solvent extraction of plant exudates see Chapter 23, under *Plant Exudates*.

Tinctures—Tinctures are defined in the USP as being alcoholic or hydroalcoholic solutions prepared from vegetable materials or from chemical substances, an example of the latter being Iodine Tincture. Traditionally, tinctures of potent vegetable drugs essentially represent the activity of 10 g of the drug in each 100 mL of tincture, the potency being adjusted following assay. Most other tinctures of vegetable drugs represent the extractive from 20 g of the drug in 100 mL of tincture.

The USP specifically describes two general processes for preparing tinctures, one by percolation designated as Process P, and the other by maceration, as Process M. These utilize the methods described under *Extraction*.

Process P includes a modification so that tinctures that require assay for adjustment to specified potency thus may be tested before dilution to final volume. A tincture prepared by Process P as modified for assayed tinctures is Belladonna Tincture.

Examples of tinctures prepared by Process M are Compound Benzoin Tincture and Sweet Orange Peel Tincture (the latter contains the extractive from 50 g of sweet orange peel in 100 mL of tincture).

Fluidextracts—The USP defines fluidextracts as being liquid preparations of vegetable drugs, containing alcohol as a solvent or as a preservative, or both, so made that each mL contains the therapeutic constituents of 1 g of the standard

drug that it represents. While the USP states that pharmacopoial fluidextracts are made by percolation, the official compendia previously have described general procedures for three percolation methods used in making fluidextracts.

Process A is a percolation method that can be modified for fluidextracts that must be assayed.

Process E is an alternative for Process A in which percolation is conducted on a column of drug much greater in length than in diameter.

Process D is used for preparing fluidextracts with boiling water as the menstruum, alcohol being added as a preservative to the concentrated percolate; this is the procedure used for preparing Cascara Sagrada Fluidextract.

The BP and PC use the designation *Liquid Extracts* for fluidextracts.

Extracts—Extracts are defined in the USP as concentrated preparations of vegetable or animal drugs obtained by removal of the active constituents of the respective drugs with suitable menstrua, evaporation of all or nearly all of the solvent and adjustment of the residual masses or powders to the prescribed standards.

Three forms of extracts are recognized: semiliquids or liquids of syrupy consistency, plastic masses (known as *pilular* or *solid extracts*) and dry powders (known as *powdered extracts*). Extracts, as concentrated forms of the drugs from which they are prepared, are used in a variety of solid or semisolid dosage forms. The USP states that pilular extracts and powdered extracts of any one drug are interchangeable medicinally, but each has its own pharmaceutical advantages. Pilular extracts, so-called because they are of a consistency to be used in pill masses and made into pills, are suited especially for use in ointments and suppositories. Powdered extracts are suited better for incorporation into a dry formulation, as in capsules, powders or tablets. Semiliquid extracts, or extracts of a syrupy consistency, may be used in the manufacture of some pharmaceutical preparations.

Most extracts are prepared by extracting the drug by percolation. The percolate is concentrated, generally by distillation under reduced pressure. The use of heat is avoided where possible because of potential injurious effect on active constituents. Powdered extracts which are made from drugs that contain inactive oily or fatty matter may have to be defatted or prepared from defatted drug. For diluents that may be used to adjust an extract to prescribed standards, see the USP.

Pure Glycyrrhiza Extract USP is an example of a pilular extract; Belladonna Extract USP and Hyoscyamus Extract PC are examples of powdered extracts (the former is prepared also as a pilular extract and the latter as a liquid extract).

References

1. Ostendorf JP: *J Soc Cosmet Chem 16:* 203, 1965.
2. Chalmers L: *Soap Perfum Cosmet 44:* Jan. 29, 1971.
3. Bruch CW: *Drug Cosmet Ind 111:* 51, Oct. 1972.
4. Contos D: *Manuf Chem Aerosol News 44 (Jun):* 35, (*Aug*): 41, (*Oct*): 34, (*Dec*): 16, 1973; *45 (Jan):* 19, 1974.
5. Hickman K *et al: Science 180:* 15, 1973.
6. Bergen A: *Can J Hosp Pharm 30 (Jul–Aug):* 109, 1977.
7. Meer T: *Flavoring Pharmaceutical Preparations* (SK&F Selected Pharm Res Refs No 4), SmithKline Corp, Philadelphia, Feb 11, 1957.
8. Schumacher GE: *Am J Hosp Pharm 24:* 688, 713, 1967; *25:* 154, 1968.
9. *Am Drug 176:* 24, May 1977.
10. Ryder J: *Int J Pharm Technol Prod Manuf 1(1):* 14, 1979.
11. *The Chemistry and Rheology of Water-Soluble Gums and Colloids*, Society of Chemical Industry Monograph, No 24, London, 1966.
12. Sumner ED: *JAPhA NS8:* 259, 1968.
13. Reynolds JEF, ed: *Martindale's, The Extra Pharmacopoeia*, 28th ed, The Pharmaceutical Press, London, 425, 1982.
14. Rettig H: *Acta Pharm Technol 24:* 143, 1978 through *Int Pharm Abstr 19:* 5096, 1982.
15. Griffin WC *et al: Drug Cosmet Ind 101 (Oct):* 41, (*Nov*): 52, 1967.
16. White RF: *Pharmaceutical Emulsions and Emulsifying Agents*, 4th ed, The Chemist and Druggist, London, 1964.
17. Florence AT, Whitehill D: *Int J Pharm 11:* 277, 1982.
18. Prince LM, ed: *Microemulsions Theory and Practice*, Academic, New York, 1977.
19. Becher P: *Emulsions: Theory and Practice*, Reinhold, New York, 267, 1965.
20. Peck GE *et al: JAPhA Sci Ed 49:* 75, 1960.
21. Lin TJ: *J Soc Cosmet Chem 29:* 117, 1978; *35:* 357, 1984.
22. Eisberg N: *Mfg Chem 53 (Jan):* 27, 1982.
23. Nash RA: *Drug Cosmet Ind 97:* 843, 1965; *98:* 39, 1966.
24. See Ref 13, 947, 948.
25. Carter SJ: *"Cooper and Gunn's Dispensing for Pharmaceutical Students,"* 12th ed, Pitman Medical, London, 214, 1975.
26. Coffman HL, Huyck CL: *Am J Hosp Pharm 20:* 132, 1963.
27. Harb NSA: *Cosmet Perfum 89 (Apr):* 67, 1974 through *Int Pharm Abstr 11:* 3304, 1974; Shapiro WB: *Cosmet Toiletries 97 (Mar):* 27, 1982 through *Int Pharm Abstr 20:* 3134, 1983.
28. Kalish J: *Drug Cosmet Ind 94:* 276, 1964.
29. Westwood N: *Pharm J 208:* 153, 1972.
30. Tortorici MP: *Am J Hosp Pharm 36:* 22, 1979.
31. Scheer AJ: *Drug Cosmet Ind 128 (Apr):* 40, (*May*): 39, (*Jun*): 52, 1981.

CHAPTER 84

Parenteral Preparations

Kenneth E Avis, DSc
Emeritus Professor, Pharmaceutics
College of Pharmacy
University of Tennessee, Memphis
The Health Science Center
Memphis, TN 38163

Dosage forms of drugs are designed to make it possible to introduce a drug into the body of a human or animal patient. Since the well-being, or even the life, of the patient may be affected, the dosage form must be designed and prepared in a manner intended to promote the safety of the patient. Concurrently, it is essential that the dosage form compliment or enhance the therapeutic effectiveness of the drug.

Parenteral (Gk, *para enteron* = beside the intestine) is the route of administration of drugs by injection under or through one or more layers of the skin or mucous membranes. Since this route circumvents these highly efficient protective barriers of the human body, exceptional purity of the dosage form must be achieved. The processes used in preparing it must embody good manufacturing practices that will produce and maintain the required quality of the product. New developments in process technology and quality control should be adopted as soon as their value and reliability have been established as a means for further improving the quality of the product.

History[1]

One of the most significant events in the beginnings of parenteral therapy was the first recorded injection of drugs into the veins of living animals, in about 1657, by the architect Sir Christopher Wren. From such a very crude beginning, the technique for intravenous injection and knowledge of the implications thereof developed slowly during the next century and a half. In 1855 Dr Alexander Wood of Edinburgh described what was probably the first subcutaneous injection of drugs for therapeutic purposes using a true hypodermic syringe.

The latter half of the 19th century brought increasing concern for safety in the administration of parenteral solutions, largely because of the work of Robert Koch and Louis Pasteur. While Charles Chamberland was developing both hot-air and steam sterilization techniques and the first bacteria-retaining filter (made of unglazed porcelain), Stanislaus Limousin was developing a suitable container, the all-glass ampul. In the middle 1920s Dr Florence Seibert provided proof that the disturbing chills and fever which often followed the intravenous injection of drugs was caused by potent products of microbial growth, pyrogens, which could be eliminated from water by distillation and from glassware by heating at elevated temperatures.

Of the recent developments that have contributed to the high quality standards currently achievable in the preparation of parenteral dosage forms, the two that have probably contributed most are the development of HEPA-filtered laminar airflow and the development of membrane microfiltration for solutions. The former made it possible to achieve ultraclean environmental conditions for processing sterile products, and the latter made it possible to remove from solutions by filtration both viable and nonviable particles of microbial size and smaller. However, many other developments in recent years have produced an impressive advance in the technology associated with the safe and reliable preparation of parenteral dosage forms. The following list identifies a few of the events which have contributed to that development.

1926—Parenterals were accepted for inclusion in the fifth edition of the *National Formulary*.
1933—The practical application of freeze-drying to clinical materials was accomplished by a team of scientists at the University of Pennsylvania.
1938—The Food, Drug and Cosmetic Act was passed by Congress, establishing the Food and Drug Administration (FDA).
1944—The sterilant ethylene oxide was discovered.
1940—The Parenteral Drug Association was organized.
1961—The concept of laminar airflow was developed by WJ Whitfield.
1962—The FDA was authorized by Congress to establish current good manufacturing practices (CGMP) regulations.
1965—Total parenteral nutrition (TPN) was developed by SJ Dudrick.
1972—The Limulus Amebocyte Lysate test for pyrogens in parenteral products was developed by JF Cooper.

Administration

Injections may be classified in five general categories:

1. Solutions ready for injection.
2. Dry, soluble products ready to be combined with a solvent just prior to use.
3. Suspensions ready for injection.
4. Dry, insoluble products ready to be combined with a vehicle just prior to use.
5. Emulsions.

These injections may be administered by such routes as intravenous, subcutaneous, intradermal, intramuscular, intraarticular and intrathecal. The nature of the product will determine the particular route of administration that may be employed. Conversely, the desired route of administration will place requirements on the formulation. For example, suspensions would not be administered directly into the blood stream because of the danger of insoluble particles blocking capillaries. Solutions to be administered subcutaneously require strict attention to tonicity adjustment, otherwise irritation of the plentiful supply of nerve endings in this anatomical area would give rise to pronounced pain. Injections intended for intraocular, intraspinal, intracisternal and intrathecal administration require the highest purity standards because of the sensitivity of nerve tissue to irritant and toxic substances.

When compared with other dosage forms, injections possess select advantages. If immediate physiological action is

needed from a drug, it usually can be provided by the intravenous injection of an aqueous solution. Modification of the formulation or another route of injection can be used to slow the onset and prolong the action of the drug. The therapeutic response of a drug is controlled more readily by parenteral administration since the irregularities of intestinal absorption are circumvented. Also, since the drug normally is administered by a professionally trained person, it confidently may be expected that the dose was actually and accurately administered. Drugs can be administered parenterally when they cannot be given orally because of the unconscious or uncooperative state of the patient, or because of inactivation or lack of absorption in the intestinal tract. Among the disadvantages of this dosage form are the requirement of asepsis at administration, the risk of tissue toxicity from local irritation, the real or psychological pain factor and the difficulty in correcting an error, should one be made. In the latter situation, unless a direct pharmacological antagonist is immediately available, correction of an error may be impossible. One other disadvantage is that daily or frequent administration poses difficulties, either for the patient to visit a professionally trained person or to learn to inject oneself.

Parenteral Combinations

Since there is a degree of discomfort for the patient with each injection, a physician frequently will seek to reduce this by combining more than one drug in one injection. This is encountered most commonly when therapeutic agents are added to large-volume solutions of electrolytes or nutrients, commonly called "IV additives," during intravenous administration. Since these are aqueous solutions, there is a high potential for chemical and physical interactions. See Chapter 85. The pharmacist is the professional best qualified to cope with these incompatibilities. However, in the past, these have been handled largely at the patient's bedside by the nurse and physician. Only recently has it been recognized that this professional area is the proper function of a pharmacist and has been so stated by the Joint Commission on Accreditation of Hospitals.[2,3]

As pharmacists have assumed increasing responsibility in this area, awareness has developed gradually of the widespread occurrence of visible, as well as invisible, physical, chemical and therapeutic incompatibilities when certain drugs are combined or added to intravenous fluids.

The development of a precipitate or a color change when preparations are combined is an immediate warning that an alteration has occurred. Such a combination should not be administered to the patient because the solid particles may occlude the blood vessels, the therapeutic agent may not be available for absorption or the drug may have been degraded into toxic substances. Moreover, in other instances, changes not visually apparent may have occurred which could be equally or more dangerous to the welfare of the patient.

The almost innumerable potential combinations present a complex situation even for the pharmacist. To aid him in making rapid decisions concerning potential problems, a number of charts have been compiled based on the visible changes that may be observed when two or more preparations are combined. However, the advent of data storage and retrieval systems using computers has provided a means to organize and gain rapid access readily to such information. The value of such information is limited by such factors as frequent changes in commercial products, variations in order of mixing or the proportions in the mixture, differences in concentration of each ingredient or variations in the period of time that the combination is held before use.

As studies have been undertaken and more information has been gained, it has been shown that knowledge of variable factors such as pH and the ionic character of the active constituents aids substantially in understanding and predicting potential incompatibilities. Kinetic studies of reaction rates may be used to describe or predict the extent of degradation. Ultimately, a thorough study should be undertaken of each therapeutic agent in combination with other drugs and IV fluids, not only of generic but of commercial preparations, from the physical, chemical and therapeutic aspects.

Ideally, no parenteral combination should be administered unless it has been studied thoroughly to determine its effect on the therapeutic value and the safety of the combination. However, such an ideal situation does not and may never exist. Therefore, it is the responsibility of the pharmacist to be as familiar as possible with the physical, chemical and therapeutic aspects of parenteral combinations and to exercise the best possible judgment as to whether or not the specific combination extemporaneously prescribed is suitable for use in a patient. A service to pharmacists has been provided through reviews of this subject.[4]

General Requirements

An inherent requirement for parenteral preparations is that they be of the very best quality and provide the maximum safety for the patient. Therefore, the pharmacist, being responsible for their preparation, must use his skills and resourcefulness at the highest level of efficiency to achieve this end. Among the areas requiring dedicated attention are the following:

1. Possession and application of high moral and professional ethics. Even the thought of using inferior techniques or ingredients in a manufacturing process must not be countenanced. The proper attitude of the person responsible for the preparation of the product is its most vital ingredient.
2. The pharmaceutical training received must be used to the fullest measure. The challenges to this knowledge bank will be many and varied.
3. Specialized techniques will be required for the manufacture of sterile preparations, employing them with alertness and sound judgment. These must be subjected to continuous critical review for faults, omissions and improvements.
4. Ingredients of the highest quality obtainable must be used. At times, ingredients may require special purification beyond that of the commercial supply. This normally will require that cost factors be given second place in importance.
5. The stability and effectiveness of the product must be established with substantiating data, either from original or published sources. This must take into account process variations and differences in ingredient specifications from one production site to another.
6. A well-defined and controlled program must be established to assure the quality of the product and the repetition of valid production procedures. This involves the evaluation of all ingredients, vigilant controls of all steps in the production procedures and careful evaluation of the finished product.

Injections or other sterile products rarely are prepared in the community pharmacy because of the lack of adequate facilities and trained personnel necessary to prepare a reliable and safe product.

In some hospital pharmacies injections are prepared from raw materials for research purposes or in the early phases of clinical studies. In most hospital pharmacies aseptic processing often is used for adding commercially available parenteral drug products to IV solutions for an individual patient. Increasingly, hospital pharmacies or independent units are dispensing parenterals for the home care of patients. Since the products dispensed most frequently are to provide the total parenteral nutrition (TPN) requirements of a patient, and these are excellent nutritional preparations for microorganisms as well, strict requirements for sterility must be met in preparation and packaging.

The preparation of the vast majority of injectable products used clinically occurs in the highly technologically advanced plants of the pharmaceutical industry. The operations are subject to the oversite of the Food and Drug Administration (FDA) through the application of the Current Good Manufacturing Practices (CGMPs) Regulations.[5] These regulations are discussed more fully in Chapter 107. While the oversite by the FDA has encouraged strongly the achieving of the essential high quality of parenterals today, the parenteral industry has taken the leadership and initiative in the extensive technological development and improvement in the quality, safety, effectiveness and administrative proficiency of parenteral dosage forms in recent years.

General Process

The preparation of a parenteral product may be considered to encompass four general areas as follows:

1. Procurement and selection of the components and containers.
2. Production facilities and procedures.
3. Control of quality.
4. Packaging and labeling.

The components of the product to be procured include vehicles, solutes, containers and closures. The steps constituting production include maintaining facilities and equipment, preparing and controlling the environment, cleaning the containers and equipment, preparing the product, filtering the solution, filling containers with the product, sealing the containers and sterilizing the product. The control of quality includes the evaluation of the components, validation of equipment and processes, determination that the production has been executed within prescribed requirements, and performance of necessary evaluative tests on the finished product. The final area of packaging and labeling includes all steps necessary to identify the finished product and enclose it in such manner that it is safely and properly prepared for sale and delivery to the user.

Components and Containers

Establishing specifications to insure the quality of each of the components of an injection is an essential first step. These specifications will be coordinated with the requirements of the specific formulation and necessarily will not be identical for a particular component if used in several different formulations.

The most stringent requirements normally will be encountered with aqueous solutions, particularly if the product is to be sterilized at an elevated temperature where reaction rates will be accelerated greatly. Modification of aqueous vehicles to include a glycol, or replacement with a nonaqueous vehicle, usually will reduce reaction rates. Dry preparations pose relatively few reaction problems but may require definitive physical specifications for ingredients that must have certain solution or dispersion characteristics when a vehicle is added.

Containers and closures are in prolonged, intimate contact with the product and may release substances or remove ingredients from the product. While not usually considered a part of a container, administration devices are a part of a container system and their effect upon the product must be assessed even though the contact period is usually brief.

Vehicles

Since most liquid injections are quite dilute, the component present in the highest proportion is the vehicle. A vehicle normally has no therapeutic activity and is nontoxic. However, it is of great importance in the formulation since it presents to body tissues the form of the active constituent for absorption. Absorption normally occurs most rapidly and completely when a drug is presented as an aqueous solution. Modification of the vehicle with water-miscible liquids or substitution with water-immiscible liquids normally decreases the rate of absorption. Absorption from a suspension may be affected by such factors as the viscosity of the vehicle, its capacity for wetting the solid particles, the solubility equilibrium produced by the vehicle and the distribution coefficient between the vehicle and aqueous body systems.

The vehicle of greatest importance for parenteral products is water. Water of suitable quality for parenteral administration must be prepared either by distillation or by reverse osmosis. Only by these means is it possible to separate adequately various liquid, gas and solid contaminating substances from water.

Preparation of Water

In general, a conventional still consists of a boiler (evaporator) containing raw water (distilland), a source of heat to vaporize the water in the evaporator, a headspace above the level of distilland with condensing surfaces for refluxing the vapor and thereby returning nonvolatile impurities to the distilland, a means for eliminating volatile impurities before the hot water vapor is condensed and a condenser for removing the heat of vaporization, thereby converting the water vapor to a liquid distillate.

The specific construction features of a still and the process specifications markedly will affect the quality of distillate obtained from a still. Those required for producing high-purity water, such as Water for Injection USP (WFI), must be considerably more stringent than those required for Purified Water USP. Among the factors that must be considered are:

1. The quality of the raw water will affect the quality of the distillate. It may be necessary that the raw water be first softened, deionized or treated by reverse osmosis to obtain a final distillate of adequate quality.
2. The size of the evaporator will affect the efficiency. It should be large enough to provide a low vapor velocity, thus reducing the entrainment of the distilland either as a film on vapor bubbles or as separate droplets.
3. The baffles (condensing surfaces) determine the effectiveness of refluxing. They should be designed to remove efficiently the entrainment at optimal vapor velocity, collecting and returning the heavier droplets contaminated with the distilland.
4. Redissolving volatile impurities in the distillate reduces its purity. Therefore, they should be separated efficiently from the hot water vapor and eliminated by aspirating them to the drain or venting them to the atmosphere.
5. Contamination of the vapor and distillate from the metal parts of the still can occur. Present standards for high-purity stills are that all parts contacted by the vapor or distillate should be constructed of metal coated with pure tin, 304 or 316 stainless steel or chemically resistant glass.

The design features of a still also influence its efficiency of operation, relative freedom from maintenance problems or extent of automatic operation. Stills may be constructed of varying size, rated according to the volume of distillate that can be produced per hour of operation under optimum conditions. Only stills designed to produce high-purity water may be considered for use in the production of WFI.

Conventional commercial stills designed for the production of high-purity water, such as shown in Fig 84-1, are available from several suppliers (*AMSCO, Barnstead, Corning, Finn-Aqua, Vaponics*).

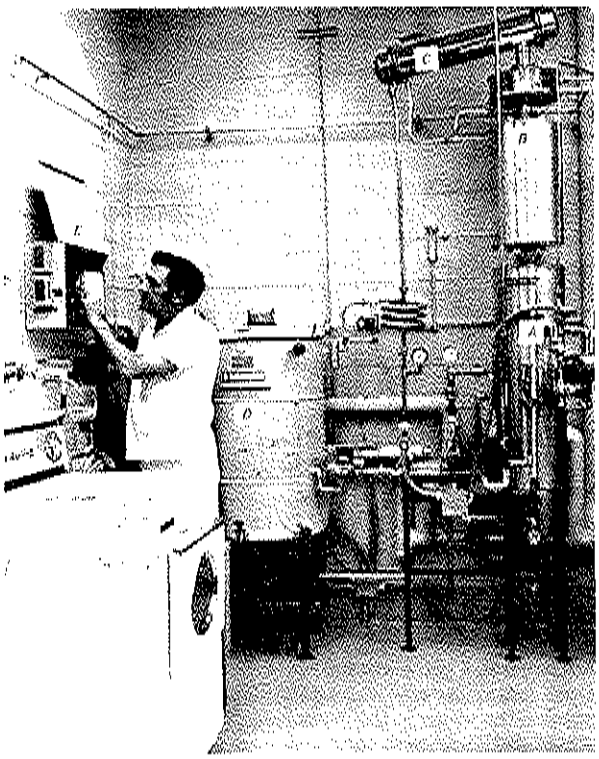


Fig 84-1. High-purity still and sealed water-storage system. *A*: evaporator; *B*: high-purity baffle unit; *C*: condenser; *D*: storage tank with ultraviolet lamp; *E*: control panel (courtesy, Ciba-Geigy).

Compression Distillation—The vapor-compression still, primarily designed for the production of large volumes of high-purity distillate with low consumption of energy and water, is illustrated diagrammatically in Fig 84-2. To start, the feed water is heated in the evaporator to boiling. The vapor produced in the tubes is separated from the entrained distilland in the separator and conveyed to a compressor which compresses the vapor and raises its temperature to approximately 224°F. It then flows to the steam chest where it condenses on the outer surfaces of the tubes containing the distilland; thereby the vapor is condensed and drawn off as a distillate while giving up its heat to bring the distilland in the tubes to the boiling point.

Vapor-compression stills are available in capacities from 50 to 2800 gal/hr (*Aqua-Chem, Barnstead, Meco*).

Multiple-Effect Stills—The multiple-effect still also is designed to conserve energy and water usage. In principle, it is simply a series of single-effect stills running at differing pressures. A series of up to seven effects may be used, with the first effect operated at the highest pressure and the last effect at atmospheric pressure. Steam from an external source is used in the first effect to generate steam under pressure from raw water; it is used as the power source to drive the second effect. The steam used to drive the second effect condenses as it gives up its heat of vaporization and forms a distillate. This process continues until the last effect when the steam is at atmospheric pressure and must be condensed in a heat exchanger.

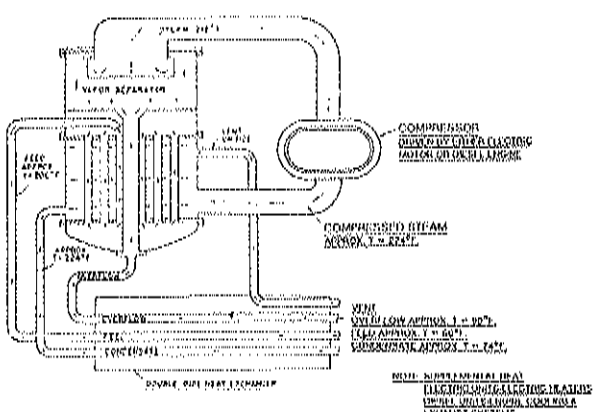


Fig 84-2. Vapor-compression still.

The capacity of a multiple-effect still can be increased by adding effects. The quality of the distillate also will be affected by the inlet steam pressure; thus, a 600-gal/hr unit designed to operate at 115 psig steam pressure could be run at approximately 55 psig and would deliver about 400 gal/hr. These stills have no moving parts and operate quietly. They are available in capacities from about 50 to 7000 gal/hr (*AMSCO, Barnstead, Finn-Aqua, Vaponics*).

Reverse Osmosis—Reverse osmosis has been added by the USP as a method suitable for preparing WFI. As the name suggests, the natural process of selective permeation of molecules through a semipermeable membrane separating two aqueous solutions of different concentrations is reversed. Pressure, usually between 200 and 400 psig, is applied to overcome osmotic pressure and force pure water to permeate through the membrane. Membranes, usually composed of cellulose esters or polyamides, are selected to provide an efficient rejection of contaminant molecules in raw water. The molecules most difficult to remove are small inorganic ones such as sodium chloride. Passage through two membranes in series is sometimes used to increase the efficiency of removal of these small molecules and to decrease the risk of structural failure of a membrane to remove other contaminants, such as bacteria and pyrogens. For additional information, see *Reverse Osmosis* in Chapter 77 (including Fig 77-14) and *Water* in Chapters 66 and 83.

Currently, extensive validation is continuing to determine whether, in fact, this method is capable of consistently producing high-purity water of a quality equal or superior to that producible by distillation. Reverse osmosis systems are available in a range of production sizes. (*AMSCO, Aqua-Chem, Finn-Aqua, Meco, Millipore*, etc).

Water for Injection USP

This is a high-purity water intended to be used as a vehicle for injectable preparations. Sterile Water for Injection USP (SWFI) is described in a separate monograph and differs in that it is intended as a packaged and sterilized product.

Storage—If WFI cannot be used immediately after it is produced, the USP permits storage at room temperature for a period not exceeding 24 hr or for longer periods at a temperature too high or too low for microbial growth to occur. Therefore, WFI usually is collected directly from the reverse-osmosis unit or a still in a closed system designed to prevent recontamination of the water and to hold it at a constant temperature of 60 to 80°C. The system may range from a relatively small single storage tank with a drawoff spout (Fig 84-1) to a very large system holding several thousand gallons of water. The stainless-steel tank in such a system usually is connected to a welded stainless-steel distribution loop supplying the various use sites with a continuously circulating water supply. The tank is provided with a hydrophobic membrane vent filter capable of excluding bacteria and nonviable particulate matter. Such a vent filter is necessary to permit changes in pressure during filling and emptying. The construction material for the tank and connecting lines is usually electropolished 316L stainless steel with heliarc welded pipe. The tanks also may be lined with

glass or a coating of pure tin. Such systems are very carefully designed and constructed and often constitute the most costly installation within the plant.

When the water cannot be used at 80°, heat exchangers must be installed to reduce the temperature at the point of use. Bacterial retentive filters should not be installed in such systems because of the risk of bacterial buildup on the filters and the consequential release of pyrogenic substances.

Purity—The USP monographs provide standards of purity for WFI and SWFI. A few of these standards require comment.

SWFI must meet the requirements of the USP Sterility Test, but WFI need not since it is to be used in a product which will be sterilized. Both must meet the requirements of the USP Pyrogen Test (page 492).

The limits for total solids varies in the two monographs. The larger the surface area of the glass container per unit volume of water, the greater the amount of glass constituents that may be leached into the water, particularly during the elevated temperature of steam sterilization.

The WFI monograph stipulates a maximum of 10 ppm of total solids. This is generally considered to be much too high to assure a quality of water that permits the stable formulation of many drugs. A relatively few metallic ions present often can render a formulation unstable. Therefore, it is common practice to set a limit of 0.1 ppm or less of ionic contaminants expressed as sodium chloride.

Ionic contaminant level is not the same as total solids; the former is a measure of only the ionic content, while the latter is a measure of the undissociated constituents as well. The ionic content of water can be measured very easily by means of a conductivity meter which frequently is used as an indicator of the purity. The results are expressed in one of three terms: as sodium chloride ions, as resistance in ohms or megohms or as conductance in micromhos. Ohms and mhos have a reciprocal relationship to each other, but they are related to ppm sodium chloride by an experimentally determined curve. To give one point of comparison, 0.1 ppm sodium chloride is equal to approximately 1.01 megohms and 0.99 micromhos. It should be mentioned that conductivity measurements give no direct indication of pyrogen content since pyrogens are undissociated organic compounds.

WFI may not contain an added substance. SWFI may contain a bacteriostatic agent when in containers of 30-mL capacity or less. This restriction is designed to prevent the administration of a large quantity of a bacteriostatic agent that probably would be toxic in the accumulated amount of a large volume of solution, even though the concentration was low.

Types of Vehicles

Aqueous Vehicles—Certain aqueous vehicles are recognized officially because of their valid use in parenterals. Often they are used as isotonic vehicles to which a drug may be added at the time of administration. The additional osmotic effect of the drug may not be enough to produce any discomfort when administered. These vehicles include Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection and Lactated Ringer's Injection.

Water-Miscible Vehicles—A number of solvents that are miscible with water have been used as a portion of the vehicle in the formulation of parenterals. These solvents are used primarily to effect the solubility of certain drugs and to reduce hydrolysis. The most important solvents in this group are ethyl alcohol, polyethylene glycol of the liquid series and propylene glycol. Ethyl alcohol is used particularly in the preparation of solutions of cardiac glycosides and the glycols in solutions of barbiturates, certain alkaloids and certain antibiotics. Such preparations usually are given intramuscularly.

These solvents, as well as nonaqueous vehicles, have been reviewed by Spiegel and Noseworthy.[6]

Nonaqueous Vehicles—The most important group of nonaqueous vehicles are the fixed oils. The USP provides specifications for such vehicles, indicating that the fixed oils must be of vegetable origin so that they will be metabolized, will be liquid at room temperature and will not become rancid readily. The USP also specifies limits for the degree of unsaturation and free fatty acid content. The oils most commonly used are corn oil, cottonseed oil, peanut oil and sesame oil. It should be noted that the official monographs for some of these oils provide for greater latitude than the specifications required for the use of the oil as a vehicle for a parenteral.

Fixed oils are used particularly as vehicles for certain hormone preparations. These and other nonaqueous vehicles, such as ethyl oleate, isopropyl myristate, and benzyl benzoate, may be used provided they are safe in the volume administered and do not interfere with the therapeutic efficacy of the preparation or with its response to prescribed assays and tests. The label also must state the name of the vehicle so that the user may beware in case of known sensitivity or other reactions to it.

Solutes

The requirements for purity of the medicinal compound used in an injection often make it necessary to undertake special purification of the usual chemical grade available. In a few instances, a special parenteral grade of a compound is available, for example, ascorbic acid freed from all traces of copper contamination. As a general rule, the best chemical grade obtainable should be used. It should be obvious that if a few ppm of ionic contaminants in WFI may cause stability problems, a similar level of contamination in the solute itself may, likewise, cause stability problems. Metallic catalysis of chemical reactions is one which is encountered frequently.

Other factors to be considered with respect to the quality of solutes include the level of microbial and pyrogenic contamination, solubility characteristics as determined by the chemical or physical form of the compound and freedom from gross dirt.

Added Substances—The USP includes in this category all substances added to a preparation to improve or safeguard its quality. An added substance may

Effect solubility, as does sodium benzoate in Caffeine and Sodium Benzoate Injection.

Provide patient comfort, as do substances added to make a solution isotonic.

Enhance the chemical stability of a solution, as do antioxidants, inert gases, chelating agents and buffers.

Preserve a preparation against the growth of microorganisms. The term "preservative" sometimes is applied only to those substances which prevent the growth of microorganisms in a preparation. However, such limited use is inappropriate, being better used for all substances that act to retard or prevent the chemical, physical or biological degradation of a preparation.

While added substances may prevent a certain reaction from taking place, they may induce others. Not only may visible incompatibilities occur, but hydrolysis, complexation, oxidation and other invisible reactions may decompose or otherwise inactivate the therapeutic agent. Therefore, added substances must be selected with due consideration and investigation of their effect on the total formulation.

Antimicrobial Agents—The USP states that antimicrobial agents in bacteriostatic or fungistatic concentrations must be added to preparations contained in multiple-dose containers. They must be present in adequate concentration at the time of use to prevent the multiplication of microorganisms inadvertently introduced into the preparation while withdrawing a portion of the contents with a hypodermic needle and syringe. Among the compounds most frequently employed, with the concentration limit prescribed by the USP, are:

Phenylmercuric nitrate and thimerosal 0.01%.
Benzethonium chloride and benzalkonium chloride 0.01%.
Phenol or cresol 0.5%.
Chlorobutanol 0.5%.

The above limit is rarely used for phenylmercuric nitrate, most frequently being employed in a concentration of 0.002%. Methyl *p*-hydroxybenzoate 0.18% and propyl *p*-hydroxybenzoate 0.02% in combination, and benzyl alcohol 2% also are used frequently. In oleaginous preparations, no antibacterial agent commonly employed appears to be effective. However, it has been reported that hexylresorcinol 0.5% and phenylmercuric benzoate 0.1% are moderately bactericidal.

Antimicrobial agents must be studied with respect to compatibility with all other components of the formula. In addition, their activity must be evaluated in the total formula. It is not uncommon to find that a particular agent will be effective in one formulation but ineffective in another. This may be due to the effect of various components of the formula on the biological activity or availability of the compound; for example, the binding and inactivation of esters of *p*-hydroxybenzoic acid by macromolecules such as Polysorbate 80 or the reduction of phenylmercuric nitrate by sulfide residues in rubber closures. A physical reaction encountered is that bacteriostatic agents sometimes are removed from solution by rubber closures.

Buffers are used primarily to stabilize a solution against the chemical degradation that might occur if the pH changes appreciably. Buffer systems employed should normally have as low a buffer capacity as feasible in order not to disturb significantly the body buffer systems when injected. In addition, the buffer range and effect on the activity of the product must be evaluated carefully. The acid salts most frequently employed as buffers are citrates, acetates and phosphates.

Antioxidants are required frequently to preserve products because of the ease with which many drugs are oxidized. Sodium bisulfite 0.1% is used most frequently. The use of sulfites has been reviewed by Schroeter[7]. Acetone sodium bisulfite, sodium formaldehyde sulfoxylate and thiourea also are used sometimes. The sodium salt of ethylenediaminetetraacetic acid has been found to enhance the activity of antioxidants in some cases, apparently by chelating metallic ions that would otherwise catalyze the oxidation reaction.

Pyrogens

Pyrogens may be anticipated contaminants in crude drugs, such as antibiotics produced by fermentation, or they may be present as unexpected and unwanted contaminants in a finished product as a result of inadvertent contamination during processing. The former must be eliminated during the purification steps of the drug. The latter can be eliminated best by preventing their introduction or development during the process. In general, the presence of pyrogens in a finished product is indicative of preparation under inadequately controlled clean conditions.

Pyrogens cause a febrile reaction in human beings. Other symptoms include chills, pains in the back and legs and malaise. While pyrogens are rarely fatal, they produce significant discomfort for the patient. On the other hand, pyrogens have been shown to induce a general nonspecific resistance to microorganisms and, on this basis, have been used therapeutically. Recent findings indicate that bacterial pyrogens, when introduced into the body, stimulate the production of an endogenous (leukocytic) pyrogen that causes the familiar physiological responses.

Pyrogens are products of the growth of microorganisms. The most potent pyrogenic substances are produced by Gram-negative bacteria (endotoxins), but Gram-positive bacteria and fungi also produce pyrogenic substances of lesser potency. Chemically endotoxins have been shown to be a phospholipid attached to a polysaccharide carrier.

Pyrogens can be destroyed by heating at high temperatures. The recommended procedure for depyrogenation of glassware and equipment is heating at a temperature of 250° for 45 min. It has been reported that 650° for 1 min or 180° for 4 hr likewise will destroy pyrogens. The usual autoclaving cycle will not do so. Heating with strong alkali or oxidizing solutions will destroy pyrogens. It has been claimed that thorough washing with detergent will render glassware pyrogen-free if protected during manufacture and storage from heavy pyrogenic contamination. Likewise, plastic containers and devices must be protected from pyrogenic contamination during manufacture and storage since known ways to destroy pyrogens will affect the plastic adversely. It has been reported that anion-exchange resins will adsorb pyrogens from water and reverse osmosis will eliminate them. However, the most reliable method for their elimination from water is distillation.

A method that has been used for the removal of pyrogens from solutions is adsorption on adsorptive agents. However, since the adsorption phenomenon also may cause selective removal of chemical substances from the solution and the filtrate may be contaminated with the agent, this method has limited application. Other in-process methods for their destruction or elimination include selective extraction procedures and careful heating with dilute alkali, dilute acid or mild oxidizing agents. In each instance, the method must be studied thoroughly to be sure it will not have an adverse effect on the constituents of the product. New developments in ultrafiltration now make possible pyrogen separation on a molecular weight basis and the process of tangential flow increasingly is making large-scale processing a reality.

Sources of Pyrogens—Pyrogens may enter a preparation by any means that will introduce living or dead microorganisms. Perhaps the greatest potential source of such contamination is the water used in processing. Although proper distillation will provide pyrogen-free water, storage conditions must be such that microorganisms are not introduced and subsequent growth is prevented.

Another potential source of contamination is equipment. Pyrogenic materials adhere strongly to glass and other surfaces. Residues of solutions in used equipment often become bacterial cultures with subsequent pyrogenic contamination. Even washed equipment left wet and exposed to the atmosphere may contain sufficient nutrients for microorganism growth. Since drying does not destroy pyrogens, they may remain in equipment for long periods. Adequate washing greatly will reduce and subsequent dry-heat treatment will render contaminated equipment suitable for use.

The solute may be a source of pyrogens. Solutes may be crystallized or precipitated from aqueous liquids containing pyrogenic contamination. In the process, pyrogens may be trapped within the particle layers. In such cases the solute must be purified by recrystallization, precipitate washing or other means of eliminating pyrogens.

The manufacturing process must be carried out with great care and as rapidly as possible to minimize the risk of microbial contamination. Preferably, no more product should be prepared than can be processed completely within one working day, including sterilization.

Containers

Containers are an integral part of the formulation of an injection and may be considered a component, for there is no container that is totally insoluble or does not in some way affect the liquid it contains, particularly if the liquid is aqueous. Therefore, the selection of a container for a particular injection must be based on a consideration of the composition of the container, as well as of the solution, and the treatment to which it will be subjected.

Table I provides a generalized comparison of the three compatibility properties—leaching, permeation and adsorption—of container materials most likely to be involved in the formulation of aqueous parenterals. Further, the integrity of the container/closure system depends upon a series of characteristics which determine the effectiveness with which it achieves its role. These considerations have been reviewed by Morton.[8]

Plastic

Thermoplastic polymers have been established as packaging materials for sterile preparations such as large-volume parenterals, ophthalmic solutions and, increasingly, for small-volume parenterals. For such use to be acceptable a thorough understanding of the characteristics, potential problems and advantages for use must be developed. One thorough review of these factors relative to pharmaceuticals has been prepared by Autian.[9] He stated that three principal problem areas exist in using these materials; namely,

1. Permeation of vapors and other molecules in either direction through the wall of the plastic container.
2. Leaching of constituents from the plastic into the product.
3. Sorption (absorption and/or adsorption) of drug molecules or ions on the plastic material.

Permeation, the most extensive problem, may be troublesome by permitting volatile constituents or selected drug molecules to migrate through the wall of the container to the outside and thereby be lost. The reverse of this also may occur by which oxygen or other molecules may permeate to the inside of the container and cause oxidative or other degradation of susceptible constituents. Leaching may be a problem when certain constituents of the plastic material migrate into the product. This potential problem often may be controlled by careful formulation of the polymer mixture with a minimum of additives. Sorption seems to be a limited problem in the packaging of parenterals and is found most commonly in association with polyamides such as nylon.

One of the principal advantages of using plastic packaging materials is that they are not breakable as is glass; also, there is a substantial weight reduction. The flexibility of the low-density polyethylene polymer, for ophthalmic preparations, makes it possible to squeeze the side wall of the container and discharge one or more drops without introducing contamination into the remainder of the product. The flexible bags of polyvinyl chloride or select polyolefins, currently in use for large-volume intravenous fluids, have the added advantage that no air interchange is required; the flexible wall simply collapses as the solution flows out of the bag.

Most plastic materials have the disadvantage that they are not as clear as glass and, therefore, inspection of the contents is impeded. In addition, many of these materials will soften or melt under the conditions of thermal sterilization. However, careful selection of the plastic used and control of the autoclave cycle has made thermal sterilization of some products possible, large-volume parenterals in particular. Eth-

Table I—Comparative Compatibility Properties of Container Materials

	Leaching		Permeation		Adsorption (selective)
	Extent[a]	Potential Leachables	Extent[a]	Potential Agents	Extent[a]
Glass					
Borosilicate	1	Alkaline earth and heavy metal oxides	0	N/A	2
Soda-lime	5	Alkaline earth and heavy metal oxides	0	N/A	2
Plastic Polymers					
Polyethylene					
Low density	2	Plasticizers, antioxidants	5	Gases, water vapor, other molecules	2
High density	1	Antioxidants	3	Gases, water vapor, other molecules	2
PVC	4	HCl, especially plasticizers, antioxidants, other stabilizers	5	Gases, especially water vapor and other molecules	2
Polyolefins	2	Antioxidants	2	Gases, water vapor, other molecules	2
Polypropylene	2	Antioxidants, lubricants	4	Gases, water vapor	1
Rubber Polymers					
Natural and related synthetic	5	Heavy metal salts, lubricants, reducing agents	3	Gases, water vapor	3
Butyl	3	Heavy metal salts, lubricants, reducing agents	1	Gases, water vapor	2
Silicone	2	Minimal	5	Gases, water vapor	1

[a] Approximate scale of 1 to 5 with "1" as the lowest.

ylene oxide or radiation sterilization may be employed for the empty container with subsequent aseptic filling. However, careful evaluation of the residues from ethylene oxide and their potential toxic effect must be undertaken.

Because of the relatively new use of plastic materials for packaging sterile preparations, considerable investigation is still required concerning potential interactions and other problems that may be encountered. For further details see Chapter 80.

Glass

Glass is employed as the container material of choice for most injections. It is composed principally of silicon dioxide with varying amounts of other oxides such as sodium, potassium, calcium, magnesium, aluminum, boron and iron. The basic structural network of glass is formed by the silicon oxide tetrahedron. Boric oxide will enter into this structure, but most of the other oxides do not. The latter are only loosely bound, are present in the network interstices and are relatively free to migrate. These migratory oxides may be leached into a solution in contact with the glass, particularly during the increased reactivity of thermal sterilization. The oxides thus dissolved may hydrolyze to raise the pH of the solution, catalyze reactions or enter into reactions. In a manner as yet uncertain, some glass compounds will be attacked by solutions and, in time, dislodge glass flakes into the solution. Disturbing reactions such as these, however, can be minimized by the proper selection of the glass composition.[10]

Types—The USP has aided in this selection by providing a classification of glass; namely,

Type I, a borosilicate glass.
Type II, a soda-lime treated glass.
Type III, a soda-lime glass.
NP, a soda-lime glass not suitable for containers for parenterals.

Type I glass is composed principally of silicon dioxide and boric oxide, with low levels of the nonnetwork-forming oxides. It is a chemically resistant glass (low leachability) also having a low thermal coefficient of expansion.

Types II and III glass compounds are composed of relatively high proportions of sodium oxide and calcium oxide. This makes the glass chemically less resistant. Both types melt at a lower temperature, are easier to mold into various shapes and have a higher thermal coefficient of expansion than Type I. While there is no one standard formulation for glass among manufacturers of these USP type categories, Type II glass usually has a lower concentration of the migratory oxides than Type III. In addition, Type II has been treated under controlled temperature and humidity conditions with sulfur dioxide to dealkalize the internal surface of the container. While it remains intact, this surface will increase substantially the chemical resistance of the glass. However, repeated exposures to sterilization and alkaline detergents will break down this dealkalized surface and expose the soda-lime compound. Therefore, Type II glass containers may be considered to be of relatively good chemical resistance for only one use.

The glass types are determined from the results of two USP tests: the Powdered Glass Test and the Water Attack Test. The latter is used only for Type II glass and is performed on the whole container, because of the dealkalized surface; the former is performed on powdered glass, which exposes internal surfaces of the glass compound. The results are based upon the amount of alkali titrated by 0.02 *N* sulfuric acid after an autoclaving cycle with the glass sample in contact with a high-purity distilled water.

Care must be used in selecting the glass type to be used for a particular injectable product. In general, Type I glass will be suitable for all products, although sulfur dioxide treatment is sometimes used for a further increase in resistance. Because cost must be considered, one of the other less expensive types may be acceptable. Type II glass may be suitable, for example, for a solution which is buffered, has a pH below 7 or is not reactive with the glass. Type III glass usually will be suitable principally for anhydrous liquids or dry substances.

Physical Characteristics—Examples of the physical shape of glass ampuls and vials are illustrated in Fig 84-3. Commercially available containers vary in size from 0.5 to 1000 mL. Sizes up to 100 mL may be obtained as ampuls and vials, and larger sizes as bottles. The latter are used mostly for intravenous and irrigating solutions. Smaller sizes are also available as cartridges. Ampuls and cartridges are drawn from glass tubing. The smaller vials may be made by molding or from tubing. Larger vials and bottles are made only by molding. Containers made by drawing tubing are generally optically clearer and have a thinner wall than molded containers (see Fig 84-3). Molded containers are uniform in external dimensions, stronger and heavier.

Easy-opening ampuls that permit the user to break off the tip at the neck constriction without the use of a file are marketed under the names Color-Break (*Kimble*) and Score-Break (*Wheaton*). An example of a modification of container design to meet a particular need is the double-chambered vial, under the name Univial (*Univial*), designed to contain a freeze-dried product in the lower and solvent in the upper chamber. Other examples are wide-mouth ampuls with flat or rounded bottoms to facilitate filling with dry materials or suspensions, and various modifications of the cartridge for use with disposable dosage units.

Glass containers must be strong enough to withstand the physical shocks of handling and shipping and the pressure differentials that develop, particularly during the autoclave sterilization cycle. They must be able to withstand the thermal shock resulting from large temperature changes during processing, for example, when the hot bottle and contents are exposed to room air at the end of the sterilization cycle. Therefore, a glass having a low coefficient of thermal expansion is necessary. The container also must be transparent to permit inspection of the contents.

Preparations which are light-sensitive must be protected by placing them in amber glass containers or by enclosing flint glass containers in opaque cartons labeled to remain on the container during the period of use. Silicone coatings are sometimes applied to containers to produce a hydrophobic surface as a means of reducing adherence of a heavy, costly

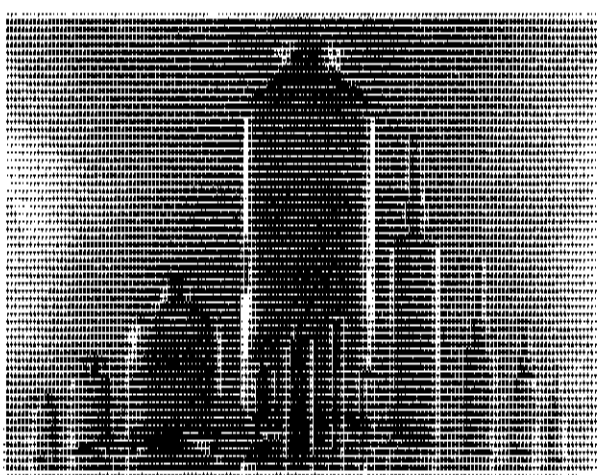


Fig 84-3. Various types of ampuls and multiple-dose vials for parenterals (courtesy, Kimble).

suspension or the friction of a rubber-tip of a syringe plunger.

The size of single-dose containers is limited to 1000 mL by the USP and multiple-dose containers to 30 mL, unless stated otherwise in a particular monograph. Multiple-dose vials are limited in size to reduce the number of punctures for withdrawing doses and the accompanying risk of contamination of the contents. As the name implies, single-dose containers are opened with aseptic care and the contents used at one time. These may range in size from 1000-mL bottles to 1-mL or less ampuls, vials or syringes. The integrity of the container is destroyed when opened so that the container cannot be closed again.

A multiple-dose container is designed so that more than one dose can be withdrawn at different times, the container maintaining a seal between uses. It should be evident that with full aseptic precautions, including sterile syringe and needle for withdrawing the dose and disinfection of the exposed surface of the closure, there is still a substantial risk of introducing contaminating microorganisms and viruses into the contents of the vial. Because of this risk, the USP requires that all multiple-dose vials must contain an antibacterial agent. However, there is no effective antiviral agent available for such use. Therefore, in spite of the advantage of flexibility of dosage provided the physician by a multiple-dose vial, the greater safety of single-dose, disposable administration units has caused their use to increase rapidly during recent years.

Rubber Closures

In order to permit introduction of a needle from a hypodermic syringe into a multiple-dose vial and provide for resealing as soon as the needle is withdrawn, each vial is sealed with a rubber closure held in place by an aluminum band. Figure 84-4 illustrates how this is done. This principle also is followed for single-dose containers of the cartridge type, except that there is only a single introduction of the needle to make possible the withdrawal or expulsion of the contents.

Rubber closures are composed of several ingredients, the primary ones being natural rubber (latex), a synthetic polymer or a combination of these. Other ingredients include a vulcanizing agent, usually sulfur; an accelerator, one of several active organic compounds such as 2-mercaptobenzothiazole; an activator, usually zinc oxide; fillers, such as carbon black or limestone and various other ingredients such as antioxidants and lubricants. These are compounded together and then vulcanized in the desired shape, making use of molds under high pressure and temperature.

Rubber closures must have sufficient elasticity to provide a snug fit between them and the lip and neck of the vial and must spring back to close the hole made by the needle immediately on withdrawal. They must not be so hard that they are highly resistant to the insertion of the needle, and they must not fragment as the hollow needle passes through them. Ideally, they should be completely nonreactive with the solution and its ingredients and should provide a complete barrier to vapor transfer. These qualities are not perfectly met by any rubber compound now available. It is, therefore, essential to determine the compatibility and performance characteristics of each rubber compound to be used.[1]

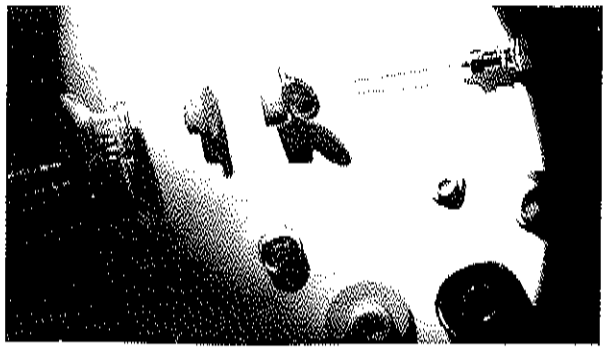


Fig 84-4. Extended view of sealing components for a multiple-dose vial (courtesy, West).

In addition to the physical tests of elasticity, hardness, fragmentation and vapor transfer, closures should be exposed to the product for prescribed periods of time at designated temperature and humidity conditions. The effect on the product of extractives from the rubber compound or loss of ingredients from the product to the closure should be determined analytically. Physicochemical and toxicological tests for evaluating rubber closures are described in the USP.

The physical shape of some typical closures may be seen in Fig 84-4. Most of them have a lip and a protruding flange that extends into the neck of the vial or bottle. Many disk closures are being used now, particularly in the high-speed packaging of antibiotics. Slotted closures are used on freeze-dried products to make it possible to insert the closure part way into the neck of the vial during the drying phase of the cycle. Partial insertion provides some protection from contamination while permitting water vapor to escape from the drying product. The plunger type is used to seal one end of a cartridge. At the time of use, the plunger expels the product by a needle inserted through the closure at the distal end of the cartridge. Intravenous solution closures often have permanent holes for adapters of administration sets; irrigating solution closures usually are designed for pouring.

Production Facilities

A product having components of the best quality quickly may become totally unacceptable if the environment in which it is processed is contaminated or if the manufacturing procedure is not carried out properly. Therefore, the production facilities and the procedure used in processing the product must meet standards adequate for the task. The nearer these standards approach perfection, the better and safer should be the product.

Arrangement of Area

The production area can be considered in terms of five functional areas: the cleanup area, the compounding area, the aseptic area, the quarantine area and the finishing or packaging area. All of these should be designed and constructed for cleaning ease, appropriate environmental control, efficient operation and personnel comfort. The extra requirements for the aseptic area are designed to provide an environment where, for example, an injection may be exposed to the environment for a brief period during subdivision from a bulk container to the individual-dose containers without becoming contaminated. Contaminants such as dust, lint and microorganisms normally are found floating in the air, lying on counters and other surfaces, on clothing and body surfaces of personnel, in the exhaled breath of personnel and deposited on the floor. The design and control of an

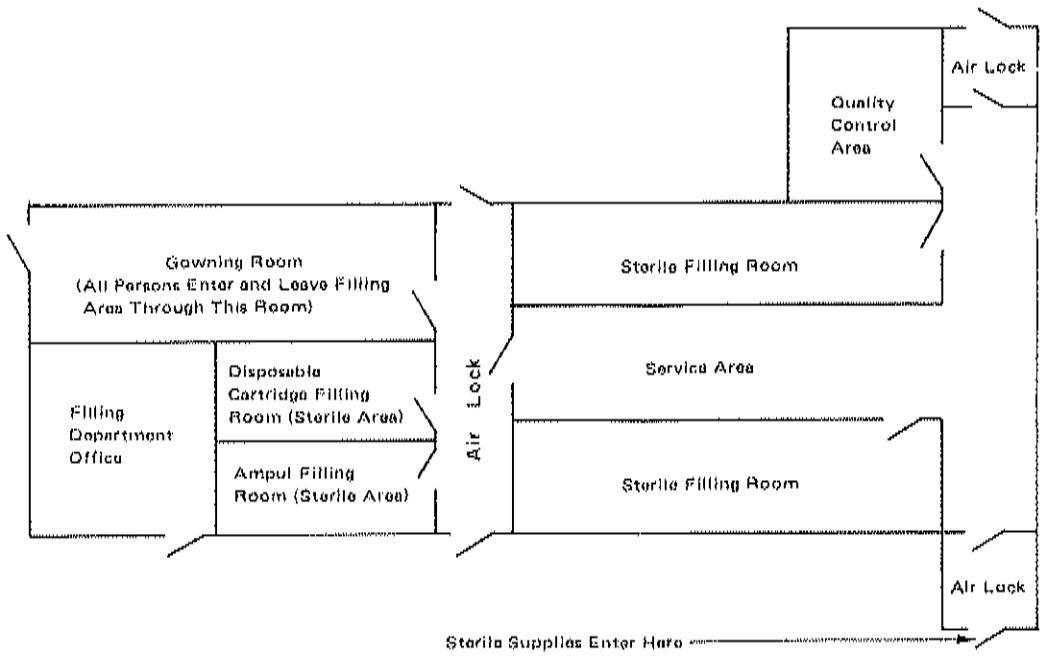


Fig 84-5. Floor plan of an aseptic filling area with its service area (courtesy, Wyeth).

aseptic area is directed toward so reducing the presence of these contaminants that they are no longer a hazard to aseptic filling. Although the aseptic area must be adjacent to support areas so that an efficient flow of components may be achieved, barriers must be provided to minimize ingress of contaminants to the aseptic area. Such a barrier may be a sealed partition, often glass-paneled for greater visibility and light. Another type of barrier is an entranceway through security doors that requires passage through an airlock so designed that both doors cannot be opened at the same time. Figure 84-5 shows an arrangement of aseptic filling rooms with adjacent support areas.

Flow Plan—In general, the components for a parenteral product flow from the stockroom, either to the compounding area, as for ingredients of the formula, or to the cleanup area, as for containers and equipment. See Fig 84-6 for a process-flow diagram. After proper processing in these areas, the components flow into the security of the aseptic area for filling of the product in appropriate containers. From there the product passes into the quarantine area where it is held until all necessary tests have been performed. If the product is to be sterilized in its final container, the passage normally is interrupted after it leaves the aseptic area for subjection to the sterilization process. After the results from all tests are known and the product has been found effective and safe, it passes to the finishing area for final labeling and packaging. There are sometimes variations from this flow plan to meet the specific needs of an individual product or to conform to available facilities. Automated operations convey the components from one area to another with little or no handling by operators.

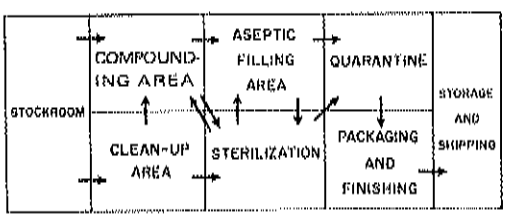


Fig 84-6. Process-flow diagram.

Cleanup Area—The cleanup area is constructed to withstand moisture, steam and detergents. The ceiling, walls, and floor should be constructed of impervious materials so that moisture will run off and not be held. One of the "spray-on-tile" finishes with a vinyl or epoxy sealing coat provides a continuous surface free from all holes or crevices. All such surfaces can be washed at regular intervals to keep them thoroughly clean. These areas should be exhausted adequately so that the heat and humidity will be removed for the comfort of personnel. Precautions must be taken to prevent the accumulation of dirt and the growth of microorganisms, especially in the presence of high humidity and heat. In this area preparation for the filling operation, such as assembling equipment, is undertaken. Adequate sink and counter space must be provided. While this area does not need to be aseptic, it must be cleanable and kept clean and the microbial load must be monitored and controlled. Precautions also must be taken to prevent deposit of particles or other contaminants on clean containers and equipment.

Compounding Area—In this area the formula is compounded. Although it is not essential that this area be aseptic, control over it should be more stringent than in the cleanup area. For example, means may need to be provided to control dust generated from weighing and compounding operations. Cabinets and counters should, preferably, be constructed of stainless steel. They should fit snugly to walls and other furniture so that there are no catch areas for dirt to accumulate. The ceiling, walls and floor should be constructed similar to those for the cleanup area. Figure 84-7 illustrates such an area located adjacent to an aseptic filling area.

Aseptic Area

This area requires construction features designed for maximum security. The ceiling, walls and floor must be sealed so that they may be washed and sanitized with a disinfectant, as needed. All counters should be constructed of stainless steel and hung from the wall so that there are no legs to accumulate dirt where they rest on the floor. All light fixtures, utility service lines and ventilation fixtures should

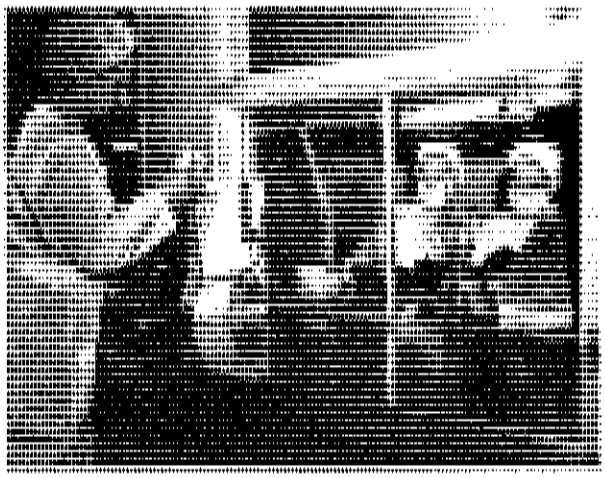


Fig 84-7. View from the service area with pipeting machine and stock-bottle retained outside of the aseptic filling area (courtesy, Wyeth).

be recessed in the walls or ceiling to eliminate ledges, joints and other locations for the accumulation of dust and dirt. As much as possible, tanks containing the compounded product should remain outside the aseptic filling area and the product fed into the area through hose lines. Figure 84-8 shows such an arrangement. Mechanical equipment that must be located in the aseptic area should be housed as completely as possible within a stainless-steel cabinet in order to seal the operating parts and their dirt-producing and accumulating tendencies from the aseptic environment. Mechanical parts that will contact the parenteral product should be demountable so that they can be sterilized.

Personnel entering the aseptic area should enter only through an airlock. They should be attired in sterile coveralls with sterile hats, masks and foot covers. Movement within the room should be minimal and in-and-out movement rigidly restricted during a filling procedure. The requirements for room preparation and the personnel may be relaxed somewhat if the product is to be sterilized in a sealed container. Some are convinced, however, that it is better to have one standard procedure meeting the most rigid requirements.

Fig 84-8. Product filtration from the aseptic staging room through a port into the aseptic filling room (courtesy, The University of Tennessee College of Pharmacy).

Air Cleaning

The air in these areas can be one of the greatest sources of contamination. It need not be, however, because several methods are available for providing clean air that is essentially free from dirt particles and microorganisms.

To provide such air, it must be cleaned thoroughly of all contaminants. This may be done by a series of treatments. Air from the outside first is passed through a prefilter, usually of glass wool, cloth or shredded plastic, to remove large particles. Then it is treated by passage through an electrostatic precipitator (Suppliers: *Am Air, Electro-Air, Sturtevant*). Such a unit induces an electrical charge on particles in the air and removes them by attraction to oppositely charged plates. The air then passes through the most efficient cleaning device, a HEPA (high efficiency particulate air) filter having an efficiency of at least 99.97% in removing particles of 0.3 μm and larger, based on the DOP (Dioctyl phthalate) test (Suppliers: *Am Air, Cambridge, Envirco, Flanders*).

For personnel comfort, air conditioning and humidity control should be incorporated into the system. Another system, the Kathabar system (*Surface Combustion*), cleans the air of dirt and microorganisms by washing it in an antiseptic solution and, at the same time, controls the humidity. The clean, aseptic air is introduced into the aseptic area and maintained under positive pressure, which prevents outside air from rushing into the aseptic area through cracks, temporarily open doors or other openings.

Laminar-Flow Environments—The required environmental control of aseptic areas has been made possible by the use of laminar-flow enclosures. Laminar airflow provides a total sweep of a confined area because the entire body of air moves with uniform velocity along parallel lines, originating through a HEPA filter occupying one entire side of the confined area. Therefore, it bathes the total area with very clean air, sweeping away contaminants.

The arrangement for the direction of airflow can be horizontal (see Fig 84-9) or vertical (see Fig 84-10), and may involve a limited area such as a workbench or an entire room. The effective air velocity is considered to be 90 ± 20 ft/min.

It must be borne in mind that any contamination introduced upstream by equipment, arms of the operator or leaks in the filter will be blown downstream. In the instance of horizontal flow this may be to the critical working site, the face of the operator or across the room. Should the contaminant be, for example, penicillin powder or viable microor-

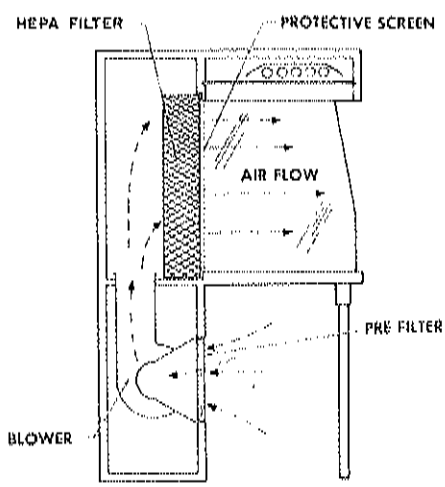


Fig 84-9. Horizontal laminar-flow workbench (courtesy, adaptation, Sandia).

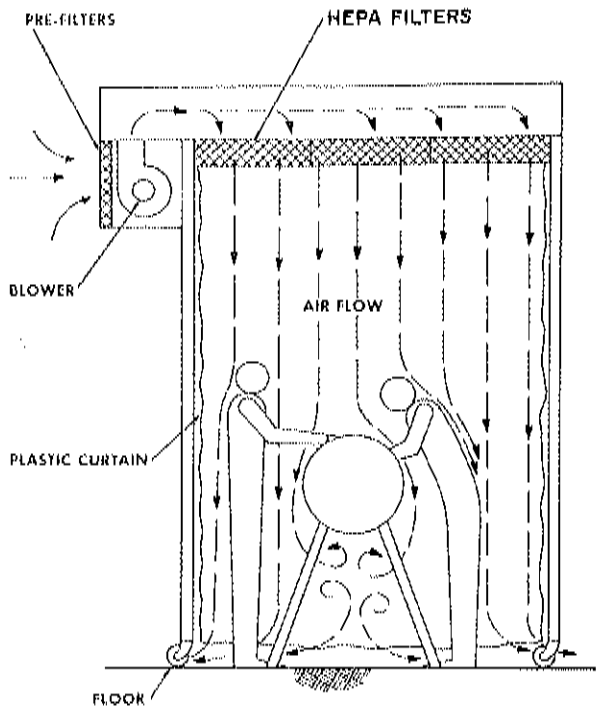


Fig 84-10. Vertical laminar-flow portable room with equipment and operators (courtesy, adaptation, Sandia).

ganisms, the danger is apparent. For operations involving such contaminants a vertical system is much more desirable, with the air flowing through perforations in the countertop or through return louvers at floor level, where it can be directed for decontamination. Vertical flow has been recommended for sterility-testing procedures.

Laminar-flow environments provide well-controlled work areas only if proper precautions are observed. Any air currents or movements exceeding the velocity of the HEPA-filtered airflow may introduce contamination, as may coughing, reaching or other manipulations of operators.

Therefore, laminar-flow work areas should be protected by being located within controlled environments. Personnel preferably should be attired for aseptic processing as described below. All movements and processes should be planned carefully to avoid the introduction of contamination upstream of the critical work area. Checks of the airstream should be performed initially and at regular intervals to be sure no leaks have developed through or around the HEPA filters. This can be done most effectively with electronic particle counters (Suppliers: *Air Techniques, Climet, Met One, Particle Measuring, Royco*).

In the manufacture of parenterals, conventional clean-room facilities frequently are supplemented by vertical laminar-airflow modules suspended above critical sites, such as filling lines. These critical operations thereby are bathed with HEPA-filtered air to provide extra protection for the product.

Laminar flow of HEPA-filtered air should meet the standard for a Class 100 clean room as defined by Federal Standard 209C,[12] which states that such an environment contains no more than 100 particles per cu ft of 0.5 μm and larger size. Conventional clean rooms would be of a lesser degree of cleanliness, such as Class 10,000, defined on the same basis. This standard has brought order into defining clean rooms and provided a common basis for their description.

Workbenches and other types of laminar-flow enclosures are available from several commercial sources (Suppliers: *Air Control, Atmos-Tech, Baker, Controlled Environment, Envirco, Flanders, Germfree, Laminaire, Liberty, Veco, Weber*).

Ultraviolet Radiation

Ultraviolet (UV) light rays have an antibacterial action, thereby producing a disinfectant action on directly irradiated surfaces. Since these rays cannot penetrate most materials, only a surface effect is produced, with the principal exception being limited penetration through air and pure water. UV light rays travel in straight lines only; therefore, objects in the path of the light beam will cast shadows with a resultant lack of irradiation in the shadow area.

UV rays are irritating to the skin and, particularly, the eyes of human beings. Therefore, personnel in the area of irradiation must be protected from direct exposure.

UV lamps may be installed so as to provide either direct or indirect radiation. Direct irradiation of a room when personnel are not present is a valuable means of reducing the bacterial count on working surfaces and floors. Lamps installed above head level, so that personnel present are not irradiated, can irradiate circulating air to reduce the microbial level continuously during processing.

Local irradiation may be useful in hood-type fixtures, over filling and other process operations, within large storage tanks or in any place where additional protection from contamination is needed, provided any product present is not affected adversely by UV rays. UV lamps usually are not employed in conjunction with laminar-flow facilities because the HEPA-filtered air sweeps exposed surfaces clean and the air itself flows too fast for adequate lethal irradiation of microorganisms being carried in the air stream.

The best practical source of UV light rays is the cold-cathode mercury vapor lamp. This lamp emits a high proportion of radiation at the 253.7 nm wavelength. A special glass is used for the tube so that the rays will pass to the outside. This glass gradually will change in crystal structure with use so that passage of the rays is gradually reduced. Such lamps, therefore, rarely burn out as do visible-light lamps but gradually reach an emission level which is ineffective. These lamps also must be kept clean, for dust and grease will lower the effective emission drastically. It generally is stated that an irradiation intensity of 20 $\mu w/cm^2$ is required for effective antibacterial activity.

Maintenance of the Aseptic Area

Important aspects in the control of environmental contamination in the aseptic area are housekeeping and maintenance. These should not be done in a haphazard manner by the general maintenance crews, but rather by crews given special instruction and under the supervision of personnel trained in the care of such areas. In general, cleaning and maintenance should be done after the completion of the day's work with an interval of quietude before the beginning of another aseptic operation. With the advent of laminar flow of HEPA-filtered air the rigors of cleaning have been reduced since the clean airflow continuously "sweeps" the area clean. All maintenance equipment should be selected for its effectiveness and freedom from lint-producing tendencies and should be reserved for use in aseptic areas only.

Personnel

Personnel selected to work on the preparation of a parenteral product must be neat, orderly and reliable. They should be in good health and free from dermatological conditions that might increase the microbial load. If they show symptoms of a head cold or similar illness, they should not

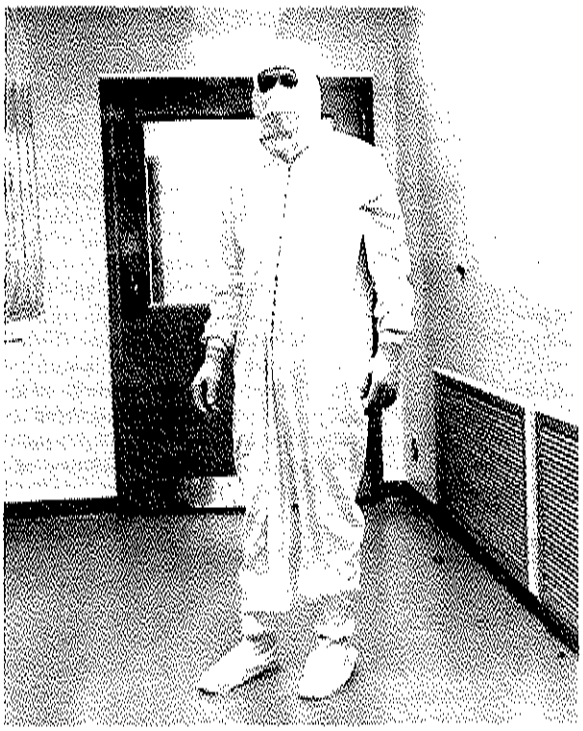


Fig 84-11. Appropriate uniform for operators entering an aseptic filling room (courtesy, Abbott).

be permitted in the aseptic area until their recovery is complete. They must receive intensive instruction in the principles of aseptic processes. They also must be made to appreciate the vital part that every movement they make has in determining the reliability and safety of the final product. Supervisors should be selected with particular care. They must be individuals who understand the unique requirements of aseptic procedures and who are able to obtain the full participation of other employees in fulfilling these exacting requirements.

The attire prescribed for personnel varies from one manufacturing facility to another. However, uniforms should be freshly laundered for each day. For use in the aseptic area, uniforms should be sterile. Usually fresh, sterile uniforms should be used after every break period, or whenever the individual returns to the aseptic area. In some plants this is not required if the product is to be sterilized in its final container. The uniform usually consists of coveralls for both men and women, hoods to completely cover the hair, face masks and Dacron or plastic boots (Fig 84-11). Sterile rubber gloves also may be required for most aseptic operations, preceded by thorough scrubbing of the hands with a disinfectant soap. In addition, goggles may be required to complete the coverage of all skin areas. The uniform is designed to confine the contaminants discharged from the body of the operator, thereby preventing their entry into the product environment.

Dacron or Tyvek uniforms are used usually, are effective barriers to discharged body particles (viable and nonviable), are essentially lint-free and are reasonably comfortable. Air showers are sometimes directed on personnel entering the processing area to blow loose lint from the uniforms.

Environmental Control Tests

In spite of the elaborate precautions taken by pharmaceutical manufacturers to provide satisfactory conditions for the proper processing of parenterals, the air may become laden with bacteria or other particles with subsequent contamination of the product. To monitor this condition, suitable environmental control tests should be performed at regular intervals.

Such tests generally are designed to measure either the particles in a volume of sampled air or the particles that are settling or have settled onto surfaces. A volume of air measured by an electronic particle counter will detect all particles and not differentiate between viable and nonviable ones. However, because of the need to control the level of microorganisms in the environment in which sterile products are processed, it also is necessary to detect viable particles, which usually are less in number than nonviable ones.

Locations for sampling should be planned to reveal potential contamination levels which may be critical in the control of the environment. For example, the most critical process step is usually the filling of dispensing containers, a site obviously requiring monitoring. Other examples include the gowning room, high-traffic sites in and out of the filling area, the penetration of conveyor lines through walls and sites near the inlet and exit of the air system.

The size of the sample should be large enough to obtain a meaningful particle count. At sites where the count is expected to be low the size of the sample may need to be increased; for example, in Class 100 areas, Whyte and Niven,[13] suggest that the sample should be at least 30 cu ft and, probably, much more. They also suggest that settling plates should be exposed in Class 100 areas for an entire fill (up to 7 to 8 hr) rather than the more common 1 hr. However, excessive dehydration of the medium must be avoided, particularly in the path of laminar-flow air.

To measure the total particle content in an air sample, electronic particle counters are available, operating on the principle of the measurement of light scattered from particles as they pass through the cell of the optical system (Suppliers: *ATI, Climet, Met One, Particle Measuring, Royco*). These instruments not only count particles but also provide a size distribution based on the magnitude of the light scattered from the particle.

Several air-sampling devices are used to obtain a count of microorganisms in a measured volume of air. A slit-to-agar (STA) sampler (*Mattson-Garvin, New Brunswick*) draws by vacuum a measured volume of air through a narrow opening causing the air to impact on the surface of a slowly rotating nutrient agar plate. Microorganisms adhere to the surface of the agar and grow into visible colonies which are counted as colony forming units (CFUs), since it is not known whether the colonies arise from a single microorganism or a cluster. A centrifugal sampler (*Biotest*) pulls air into the sampler by means of a rotating propeller and slings the air by centrifugal action against a peripheral nutrient agar strip. The advantages of this unit are that it can be disinfected easily and is portable so that it can be hand-carried wherever needed. These two methods are used quite widely.

Another volumetric air sampler is an open-faced filter holder (*Gelman, Millipore, Nuclepore, Sartorius*). The air sample is drawn through the filter membrane in the holder by means of a vacuum, the volume being controlled by means of a limiting orifice. This device can be used for obtaining either a total particle count or a count of CFUs, depending on whether the membrane is subsequently placed on a microscope slide and examined under the microscope for particles or placed on nutrient agar medium and incubated for the growth of CFUs.

It should be noted that most vegetative forms of microorganisms will be dehydrated and killed by the dehydrating effect of the airstream; therefore, the CFUs would arise principally from the growth of spores. Another device is the liquid impinger. An air sample is drawn into the orifice of

the sampler by vacuum through a limiting orifice and bubbled through a dilute nutrient medium or saline. The objective is to wash microorganisms out of the air bubbles and into the liquid medium which then is filtered through a membrane filter, the membrane placed on the nutrient agar medium and incubated. This method is somewhat more complex, but it is used in aerobiology as a reference method. Vegetative microorganisms are likely to survive because of the relatively soft impingement in the liquid.

A widely used method for microbiological sampling consists of the exposure of nutrient agar culture plates to the settling of microorganisms from the air. This method is very simple and inexpensive to perform but will detect only those organisms which have settled on the plate; therefore, it does not measure the number of microorganisms in a measured volume of air. Nevertheless, if the conditions of exposure are repeated consistently, a comparison of CFUs at one sampling site from one time to another can be meaningful.

The level of microorganisms on surfaces can be determined with nutrient agar plates having a convex surface (*Rodac Plates*). With these it is possible to roll the raised agar surface over flat or irregular surfaces to be tested. Organisms will be picked up on the agar and will grow during subsequent incubation.

Results from the above tests are very valuable to keep cleaning, production and quality-control personnel apprised of the level of contamination in a given area and, by comparison with baseline counts, will indicate when more extensive cleaning and sanitizing is needed. The results also may serve to detect environmental control defects such as failure in air-cleaning equipment or the presence of personnel who may be disseminating large numbers of bacteria without apparent physical ill effects.

Another test which is much more stringent is the filling and sealing of sterile trypticase soy broth in sterile containers under the same conditions used for an aseptic fill of a product, a "media fill." The entire lot then is incubated and examined subsequently for the appearance of growth of microorganisms which is indicative of contamination from the environment, the process, the operators or the equipment. It also may be used as a measure of the efficiency of a particular operator. Since this is a "total sterility test," it is the best indication of the efficiency of the aseptic filling process.

Production Procedures

Cleaning Containers and Equipment

Containers and equipment coming in contact with parenteral preparations must be cleaned meticulously. It is obvious that if this were not so, all other precautions to prevent contamination of the product would be useless. It also should be obvious that even new, unused containers and equipment will be contaminated with such debris as dust, fibers, chemical films and other materials arising from such sources as the atmosphere, cartons, the manufacturing process and human hands. Much greater contamination must be removed from previously used containers and equipment before they will be suitable for reuse. Equipment should be reserved rigidly for use only with parenteral preparations and, where conditions dictate, only for one type of product in order to reduce the risk of contamination.

A variety of machines are available for cleaning containers for parenteral products. These vary in complexity from a single-jet tube for hand rinsing one inverted container at a time with distilled water, to complex, automatic washers capable of processing several thousand containers an hour. The selection of the particular type will be determined largely by the physical type of containers, their condition with respect to contamination, and the number to be processed in a given period of time.

Characteristics of Machinery—Regardless of the type of cleaning machine selected, certain fundamental characteristics usually are required.

1. The liquid or air treatment must be introduced in such a manner that it will strike the bottom of the inside of the inverted container, spread in all directions and smoothly flow down the walls and out the opening with a sweeping action. The pressure of the jet stream should be such that there is minimal splashing, and the flow should be such that it can leave the container opening without accumulating and producing turbulence inside. Splashing may prevent cleaning all areas, and turbulence may redeposit loosened debris. Therefore, direct introduction of the jet stream within the container with control of its flow is required.
2. The container must receive a concurrent outside rinse.
3. The cycle of treatment should provide for a planned sequence alternating very hot and cool treatments. The final treatment should be an effective rinse with water of a quality equivalent to WFI.
4. All metal parts coming in contact with the containers and with the treatments should be constructed of stainless steel or some other noncorroding and noncontaminating material.

Treatment Cycle—The cycle of treatments to be employed will vary with the condition of the containers to be cleaned. In general, loose dirt can be removed by vigorous rinsing with water. Detergents rarely are used for new containers because of the risk of leaving detergent residues. However, a thermal-shock sequence in the cycle usually is employed to aid, by expansion and contraction, loosening of debris that may be adhering to the container wall. Sometimes only an air rinse is used for new containers, particularly if used for a dry powder. In all instances the final rinse, whether air or WFI, must be ultraclean so that no particulate residues are left by the rinsing agent.

Containers previously used cannot be reliably cleaned and the cost of attempting to do so is prohibitive. Therefore, normally, only new containers are used for parenterals. Improvements have been made in maintaining their cleanliness during shipment from the manufacturer through tight, low-shedding packaging, including plastic blister packs.

Machinery for Containers—The machinery available for cleaning large numbers of containers embodies the above principles but varies in the mechanics by which it is accomplished. In one approach, the jet tubes are arranged on arms like the spokes of a wheel, which rotate around a center post

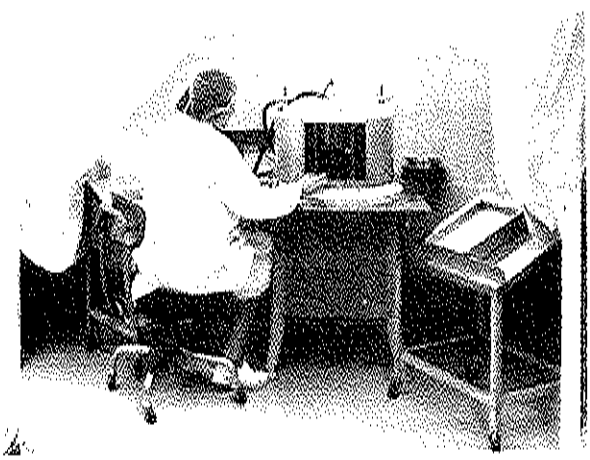


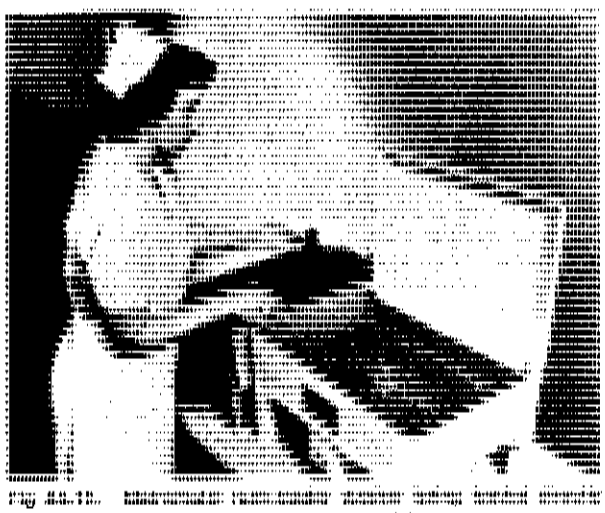
Fig 84-12. Rotary rinser (Cozzoli) in a clean environment provided by vertical laminar airflow within a curtained enclosure (courtesy, Ciba-Geigy).

through which the treatments are introduced. An operator places the unclean containers on the jet tubes as they pass the loading point and removes the clean containers as they complete one rotation. Such a machine is pictured in Fig 84-12. Another machine has a row of jet tubes across a conveyor belt. The belt moves the row of containers past the treatment stations and discharges the clean ones on the opposite end of the machine, preferably through a wall into a clean room. Two operators are required for this machine (Fig 84-13). A cabinet-type washer permits loading the containers on a rack of jet tubes. The rack is pushed inside the cabinet during the cleaning cycle. This type of machine (Fig 84-14) permits handling a variety of sizes and types of containers quite easily, but the number of containers handled in a given period of time is relatively small.

The disadvantage common to all of the above types of machines is that they require the individual handling of each container for loading and unloading. A type which overcomes this disadvantage is the rack-loading washer. Racks are prepared to fit over the open ends of ampuls or vials as they are found in shipping cartons. Inverting the carton permits the containers to be transferred from the carton to the washer without handling them individually. A battery of jet tubes is arranged to enter each container positioned in the rack. The clean containers may be removed in the rack and transferred to a box for dry-heat sterilization and storage (see Fig 84-15). More details of the industrial washing of glassware have been given by Anschel.[14]

Handling after Cleaning.—The wet, clean containers must be handled in such a way that contamination will not be reintroduced. A wet surface much more readily will collect contaminants than will a dry surface. For this reason wet, rinsed containers should be protected, such as by a laminar flow of clean air until covered, as within a stainless-steel box (see Figs 84-12 and 84-16). In addition, microorganisms are more likely to grow in the presence of moisture. Therefore, it is preferable, if not required, that containers be dry-heat sterilized in a stainless-steel box that will protect them from contamination during storage after sterilization. Doubling the heating period generally has been considered to be adequate also to destroy pyrogens, but the actual time-temperature conditions required must be validated. If it is

Fig 84-13. Conveyor rinser (Cozzoli) discharging clean vials in a preparation area (courtesy, Schering).



from a container carton (courtesy, Metromatic).

Fig 84-14. Cabinet washer (Better Built) being loaded with ampuls (courtesy, The University of Tennessee College of Pharmacy).

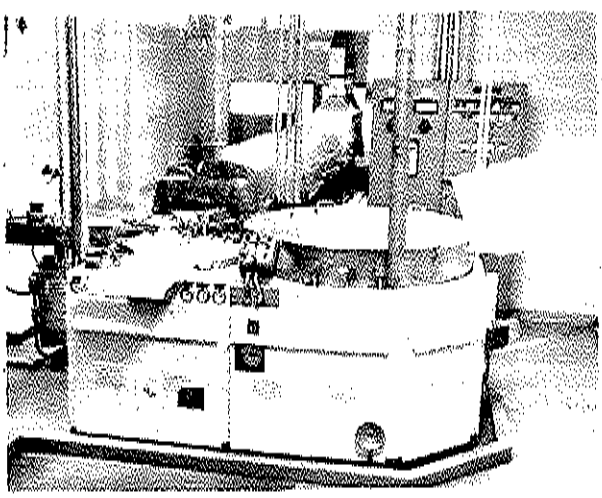


Fig 84-16. Continuous automatic line operation for vials from a rotary rinser through a sterilizing tunnel with vertical laminar-airflow protection of clean vials (courtesy, Abbott).

proved that sterilization is not essential, the containers preferably should be filled immediately with product.

Increases in process rates have necessitated the development of continuous-line processing with a minimum of individual handling, still maintaining adequate control of the cleaning and handling of the containers. Fig 84-16 shows a continuous automatic-line operation from feeding the unwashed container into the rotary rinser to passing it through the drying and sterilizing tunnel. The clean, wet containers are protected by filtered laminar-flow air from the rinser through the tunnel and until they are delivered to the filling line.

Closures—Rubber closures are coated with lubricant from the molding operation. In addition, the rough surface and electrostatic attraction tend to hold debris. Also, the surface "bloom" from migrated inorganic constituents of the compound must be removed. The recommended procedure calls for gentle agitation in a hot solution of a water softener such as 0.5% sodium pyrophosphate. The closures are removed from the solution and rinsed several times, or continuously for a prolonged period, with water and finally with filtered WFI. The rinsing is to be done in a manner which will flush away loosened debris. The wet closures then are sterilized, usually by autoclaving, and stored in closed containers until ready for use. At times this step is carried out in a solution of the bacteriostatic agent to be used in the product, in order to equilibrate the rubber closure with the agent. Subsequent loss of the agent from the solution to the closure is then less likely to occur. If the closures were immersed during autoclaving, the solution is drained off before storage to reduce hydration of the rubber compound. If the closures must be dry for use, they may be subjected to vacuum drying at a temperature in the vicinity of 100°.

The equipment used for washing large numbers of closures is usually an agitator or horizontal basket-type automatic washing machine. Because of particulate generation from the abrading action of these machines, some heat the closures in kettles in detergent solution and follow with prolonged flush rinsing. The final rinse always should be ultraclean WFI.

Equipment—The details of certain prescribed techniques for cleaning and preparing equipment, as well as of containers and closures, have been presented elsewhere.[15] Here, a few points will be emphasized.

All equipment should be disassembled as much as possible to provide access to internal structures. For thorough cleaning, surfaces should be scrubbed thoroughly with a stiff brush using an effective detergent, paying particular attention to joints, crevices, screw threads, and other structures where debris is apt to collect. Exposure to a stream of clean steam will aid in dislodging residues from the walls of stationary tanks, spigots, pipes and similar structures. Thorough rinsing with distilled water should follow the cleaning steps. Large stationary tanks, such as those shown in Fig 84-17, should be protected as much as possible from contamination after cleaning but should be rinsed thoroughly again with distilled water prior to reuse.

A relatively new concept for cleaning tanks, piping and associated attachments is called cleaning in place (CIP). Such an approach involves designing the system, normally of stainless steel, with smooth, rounded internal surfaces and without crevices. That is, for example, with welded rather than threaded connections. The cleaning is accomplished with the scrubbing action of high-pressure spray balls or nozzles delivering hot detergent solution from tanks captive to the system. Thorough rinsing with WFI follows and is accomplished within the same system. Such a process is often automated and may be computer-controlled.[16]

Rubber tubing, rubber gaskets and other rubber parts may be washed in a manner such as described for rubber closures. Thorough rinsing of tubing must be done by passing distilled water through it. However, due to the relatively porous nature of rubber compounds and the difficulty in removing all traces of chemicals from previous use, it is considered by some inadvisable to reuse rubber tubing. Rubber tubing must be left wet when preparing for sterilization by autoclaving.

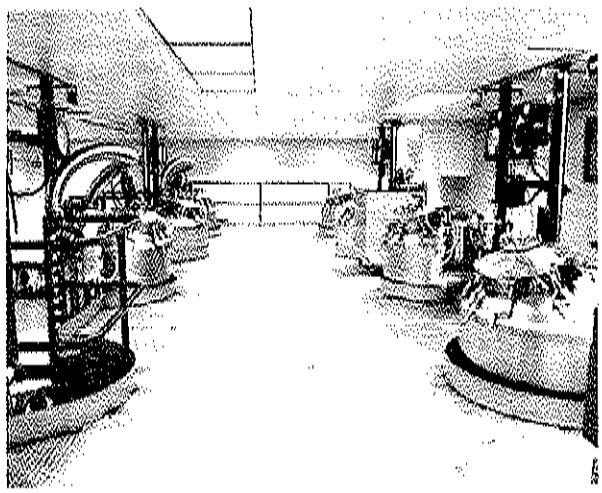


Fig 84-17. Large stainless-steel tanks for product preparation showing mezzanine access level (courtesy, Abbott).

Product Preparation

The basic principles employed in the compounding of the product do not vary from those used routinely by qualified pharmacists. However, selected aspects will be mentioned for emphasis.

All measurements should be made as accurately as possible and should be checked by a second qualified person. Although most liquid preparations are made by volume, where possible they should be made by weight, with the weight experimentally determined from a prescribed volume. This method is more accurate since no consideration need be given to the temperature of the components. In addition, measurements by weight normally can be performed more accurately than those by volume.

Care must be taken that equipment is not wet enough to significantly dilute the product or, in the case of anhydrous products, to cause a physical incompatibility. The order of mixing of ingredients may affect the product significantly, particularly those of large volume where attaining homogeneity requires considerable mixing time. For example, the adjustment of pH by the addition of an acid, even though diluted, may cause excessive local reduction in the pH of the product so that adverse effects are produced before the acid can be dispersed throughout the entire volume of product.

Parenteral dispersions, including colloids, emulsions and suspensions, provide particular problems. Parenteral emulsions have been reviewed by Singh and Ravin.[17] In addition to the problems of achieving and maintaining proper reduction in particle size under aseptic conditions, the dispersion must be kept in a uniform state of suspension throughout the preparative, transfer and subdividing operations.

The formulation of a stable product is of paramount importance. Certain aspects of this have been mentioned in the discussion of components of the product. Exhaustive coverage of the topic is not possible within the limits of this text, but further coverage is provided in Chapter 75. It should be mentioned here, however, that thermal steriliza-

tion of parenteral products increases the possibility of chemical reactions. Such reactions may progress to completion during the period of elevated temperature in the autoclave, or be initiated at this time but continue during subsequent storage. The assurance of attaining product stability requires a high order of pharmaceutical knowledge and responsibility.

Filtration

After a product has been compounded, it must be filtered if it is a solution. The primary objective of filtration is to clarify a solution. A high degree of clarification is termed "polishing" a solution. This term is used when particulate matter down to approximately 2 μm in size is removed. A further step, removing particulate matter down to 0.2 μm in size, would eliminate microorganisms and would accomplish "cold" sterilization. A solution having a high degree of clarity conveys the impression of high quality and purity, desirable characteristics for a parenteral solution.

Filters are thought to function by one or, usually, a combination of the following: (1) sieving or screening, (2) entrapment or impaction and (3) electrostatic attraction. When a filter retains particles by sieving, they are retained on the surface of the filter. Entrapment occurs when a particle, smaller than the dimensions of the passageway (pore), becomes lodged in a turn or impacted on the surface of the passageway. Electrostatic attraction causes particles opposite in charge to that of the surface of the filter pore to be held or adsorbed to the surface. It should be noted that increasing, prolonging or varying the force behind the solution may tend to sweep particles initially held by entrapment or electrostatic charge through the pores and into the filtrate.

Today, membrane filters are used almost exclusively for parenteral solutions. Their particle retention effectiveness, flow rate, nonreactivity and disposable characteristics have justified their use to the exclusion of most other types. The most common membranes are composed of:

Cellulose ester (Suppliers: *Cuno, Gelman, Millipore, Sartorius, Schleicher*).
Nylon (Supplier: *Pall*).
Polysulfone (Suppliers: *Gelman, Millipore*).
Polycarbonate (Supplier: *Nuclepore*).

but other materials are being used, including Teflon and other plastic polymers. They are available as flat membranes or pleated into cylinders to increase surface area and, thus, flow rate. Each filter in its holder should be tested for integrity before and after use, particularly if it is being used to eliminate microorganisms. While membrane filters are disposable, and thus discarded after use, the holders must be cleaned thoroughly between uses. Increasingly, clean, sterile, pretested, disposable assemblies for small as well as relatively large volumes of solutions are becoming available commercially. Other characteristics of these filters, important for a full understanding of their use, are given in Chapter 78.

Filling

During the filling of containers with a product, the most stringent requirements must be exercised to prevent contamination, particularly if the product has been sterilized by filtration and will not be sterilized in the final container. Under the latter conditions the process usually is called an "aseptic fill." During the filling operation, the product must be transferred from a bulk container and subdivided into dose containers. This operation exposes the product to the environment, equipment and manipulative technique of the operator until it can be sealed in the dose container. Therefore, this operation is carried out in the aseptic filling area where maximum protection is provided. Additional protection may be provided by filling under a blanket of HEPA-filtered laminar-flow air within the aseptic area.

Normally, the compounded product is in the form of either a liquid or a solid. A liquid is more readily subdivided uniformly and introduced into a container having a narrow mouth than is a solid. Mobile, nonsticking liquids are considerably easier to transfer and subdivide than viscous, sticky liquids, which require heavy-duty machinery for rapid production filling.

Although many devices are available for filling containers with liquids, certain characteristics are fundamental to them all. A means is provided for repetitively forcing a measured volume of the liquid through the orifice of a delivery tube which is introduced into the container. The size of the delivery tube will vary from that of about a 20-gauge hypodermic needle to a tube ½ in. or more in diameter. The size required is determined by the physical characteristics of the liquid, the desired delivery speed and the inside diameter of the neck of the container. The tube must enter the neck and deliver the liquid well into the neck to eliminate spillage, allowing sufficient clearance to permit air to leave the container as the liquid enters. The delivery tube should be as large in diameter as possible in order to reduce the resistance to the flow of the liquid. For smaller volumes of liquids, the delivery usually is obtained from the stroke of the plunger of a syringe, forcing the liquid through a two-way valve providing for alternate filling of the syringe and delivery of mobile liquids. A sliding piston valve would be used for heavy, viscous liquids. Other mechanisms include the turn of an auger in the neck of a funnel or the oscillation of a rubber diaphragm. For large volumes the quantity delivered usually is measured in the container by the level of fill in the container, the force required to transfer the liquid being provided by gravity, a pressure pump or a vacuum pump.

The narrow neck of an ampul limits the clearance possible between the delivery tube and the inside of the neck. Since a drop of liquid normally hangs at the tip of the delivery tube after a delivery, the neck of an ampul will be wet as the delivery tube is withdrawn, unless the drop is retracted. Therefore, filling machines should have a mechanism by which this drop can be drawn back into the lumen of the tube.

Since the liquid will be in intimate contact with the parts of the machine through which it flows, these must be constructed of nonreactive materials such as borosilicate glass or stainless steel. In addition, they should easily be demountable for cleaning and sterilization.

Because of the concern for particulate matter in injectable preparations, a final filter often is inserted in the system between the filler and the delivery tube. Most frequently this is a membrane filter, having a porosity of approximately 1 μm and treated to have a hydrophobic edge. This is necessary to reduce the risk of rupture of the membrane due to filling pulsations. It should be noted that the insertion of the filter at this point should collect all particulate matter generated during the process. Only that which may be found in inadequately cleaned containers or picked up from exposure to the environment after passage through the final filter potentially remain as contaminants. However, the filter does cushion liquid flow and reduces the efficiency of drop retraction from the end of the delivery tube, sometimes making it difficult to control delivery volume as precisely as would be possible without the filter.

Liquids—The filling of a small number of containers may be accomplished with a hypodermic syringe and needle, the liquid being drawn into the syringe and forced through the needle into the container. A device for providing greater speed of filling is the Cornwall Pipet (*BD & Co*). This has a

two-way valve between the syringe and the needle and a means for setting the stroke of the syringe so that the same volume will be delivered each time. Clean, sterile, disposable assemblies (Suppliers: *Burron, Pharmaseal*) operating on the same principle have particular usefulness in hospital pharmacy operations.

Mechanically operated instruments substitute a motor for the operator's hand in the previous devices described. Thereby, a much faster filling rate can be achieved. By careful engineering, the stroke of the syringe can be repeated precisely, and so, once a particular setting has been calibrated to the delivery, high delivery precision is possible. However, the speed of delivery, the expansion of the rubber tubing connecting the valve with the delivery tube, and the rapidity of action of the valves can affect the precision of delivery. A filling machine employing a two-way valve assembly is shown in operation in Fig 84-7. One employing a piston valve is shown in Fig 84-18. Stainless-steel syringes are required with viscous liquids because glass syringes are not strong enough to withstand the high pressures developed during delivery.

When high-speed filling rates are desired but accuracy and precision must be maintained, multiple filling units are often joined together in an electronically coordinated machine, such as shown in Fig 84-19.

Most high-speed fillers for large-volume solutions use the bottle as the measuring device, transferring the liquid either by vacuum or positive pressure from the bulk reservoir to the individual unit containers. Therefore, a high accuracy of fill is not achievable.

The USP indicates that each container should be filled with a slight excess of volume and gives a table of such suggested excess.

Solids—Sterile solids, such as antibiotics, are more difficult to subdivide evenly into containers than are liquids. The rate of flow of solid material is slow and irregular. Even though a container with a larger diameter opening is used to facilitate filling, it is difficult to introduce the solid particles, and the risk of spillage is ever-present. The accuracy of the quantity delivered cannot be controlled as well as with liquids. Because of these factors, the tolerances permitted for the content of such containers must be relatively large. Suggested tolerances can be found in the USP.

Some sterile solids are subdivided into containers by individual weighing. A scoop usually is provided to aid in approximating the quantity required, but the quantity filled into the container finally is weighed on a balance. This is a slow process. When the solid is obtainable in a granular form so that it will flow more freely, other methods of filling may be employed. In general, these involve the measurement and delivery of a volume of the granular material which has been calibrated in terms of the weight desired. In the machine shown in Fig 84-20 an adjustable cavity in the rim of a wheel is filled by vacuum and the contents held by vacuum until the cavity is inverted over the container. The solid material then is discharged into the container by the use of sterile air. Another machine employs an auger in the stem of a funnel at the bottom of a hopper. The granular material is placed in the hopper. By controlling the size of the auger and its rotation, a regulated volume of granular material can be delivered from the funnel stem into the container. Such a machine is shown in Fig 84-21.

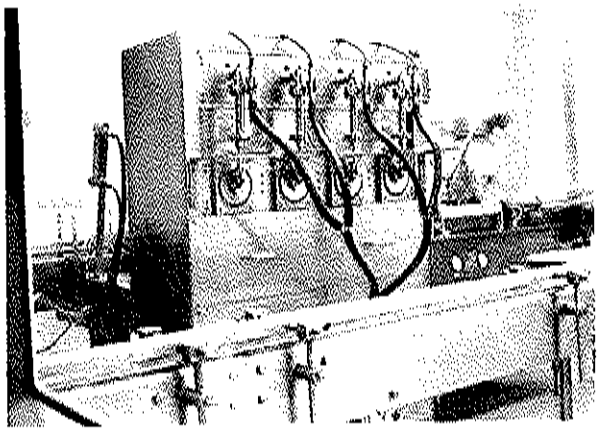


Fig 84-19. Four-pump liquid filler, with a conveyor line for vials protected by a vertical laminar airflow and plastic curtain; note the automatic stoppering machine on the right within the curtain (courtesy, Abbott).

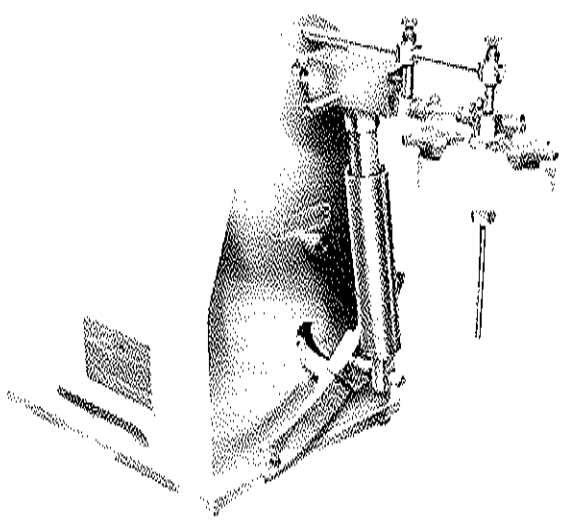


Fig 84-18. Filling machine employing a piston valve and a stainless-steel syringe (courtesy, Cozzoli).

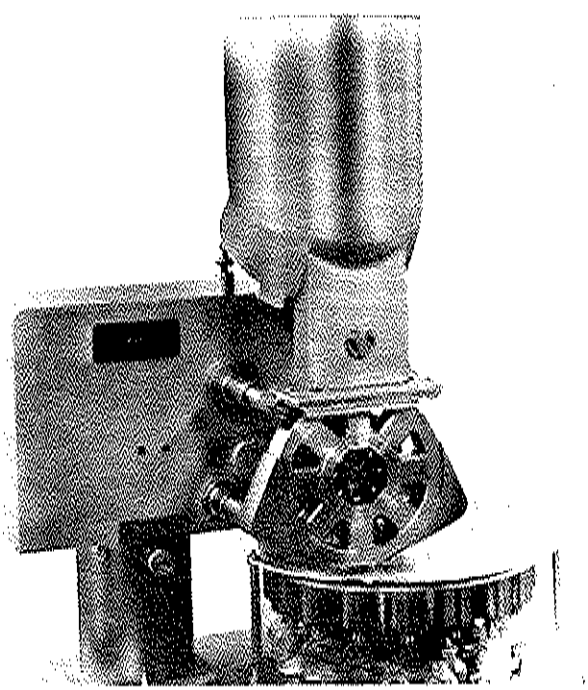


Fig 84-20. Accofil vacuum powder filler (courtesy, Perry).

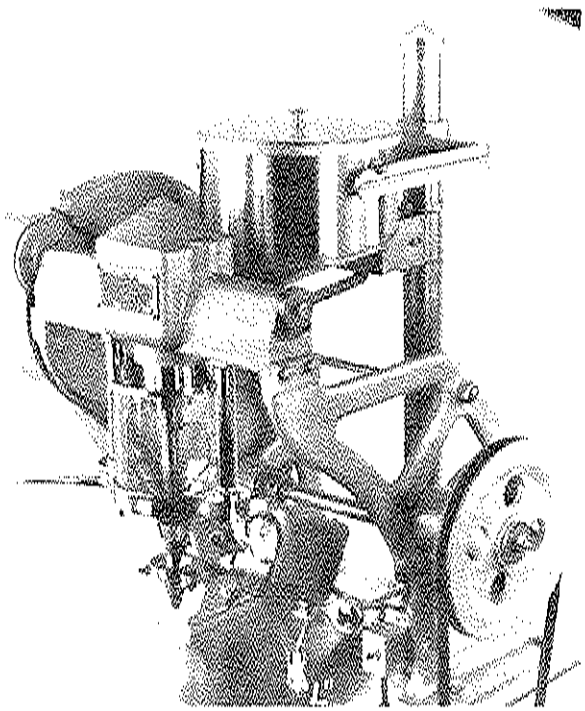


Fig 84-21. Auger-type powder filler (courtesy, Chase-Logeman).

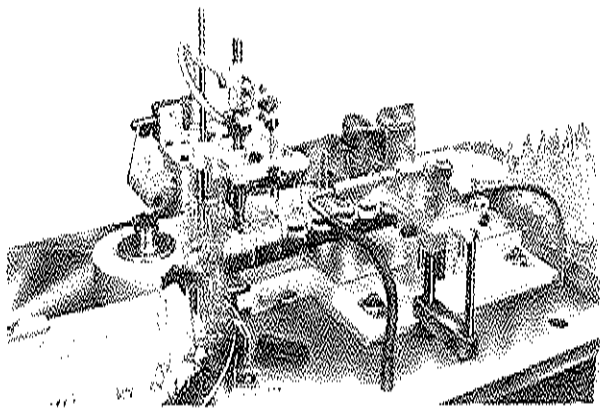


Fig 84-22. Automatic filling and pull-sealing of ampuls (courtesy, Cozzoli).

Sealing

Ampuls—Filled containers should be sealed as soon as possible to prevent the contents from being contaminated by the environment. Ampuls are sealed by melting a portion of the glass neck. Two types of seals are employed normally: tip-seals (bead-seals) or pull-seals.

Tip-seals are made by melting enough glass at the tip of the neck of an ampul to form a bead and close the opening. These can be made rapidly in a high-temperature gas–oxygen flame. To produce a uniform bead, the ampul neck must be heated evenly on all sides. This may be accomplished by means of burners on opposite sides of stationary ampuls or by rotating the ampul in a single flame. Care must be taken to properly adjust the flame temperature and the interval of heating to obtain complete closing of the opening with a bead of glass. Excessive heating will result in the expansion of the gases within the ampul against the soft bead seal and cause a bubble to form. If it bursts, the ampul is no longer sealed; if it does not, the wall of the bubble will be thin and fragile. Insufficient heating will leave an open capillary through the center of the bead. An incompletely sealed ampul is called a "leaker."

Pull-seals are made by heating the neck of the ampul below the tip, leaving enough of the tip for grasping with forceps or other mechanical devices. The ampul is rotated in the flame from a single burner. When the glass has softened, the tip is grasped firmly and pulled quickly away from the body of the ampul, which continues to rotate. The small capillary tube thus formed is twisted closed. Pull-sealing is slower, but the seals are more sure than tip-sealing. Fig 84-22 shows a machine combining the steps of filling and pull-sealing ampuls.

Powder ampuls or other types having a wide opening must be sealed by pull-sealing. Were these sealed by tip-sealing, the very large bead produced would induce glass strain with subsequent fracture at the juncture of the bead and neck wall. Fracture of the neck of ampuls during sealing also may occur if wetting of the necks occurred at the time of filling. Also, wet necks increase the frequency of bubble formation. If the product in the ampul is organic in nature, wet necks also will result in unsightly carbon deposits from the heat of sealing.

In order to prevent decomposition of a product, it is sometimes necessary to displace the air in the space above the product in the ampul with an inert gas. This is done by introducing a stream of the gas, such as nitrogen or carbon dioxide, during or after filling with the product. Immediately thereafter the ampul is sealed before the gas can diffuse to the outside.

Vials and Bottles—These are sealed by closing the opening with a rubber closure (stopper). This must be accomplished as rapidly as possible after filling and with reasoned care to prevent contamination of the contents. The large opening makes the introduction of contamination much easier than with ampuls. Therefore, a covering should be provided for such containers except for the minimal time required for filling and for the actual introduction of the rubber closure. During the latter critical time the open containers should be protected from the ingress of contamination, preferably with a blanket of HEPA-filtered laminar airflow. In Fig 84-19 the automatic conveyorized procedure is being performed under vertical laminar airflow within plastic side curtains.

The closure must fit the mouth of the container snugly enough so that its elasticity will permit adjustment to slight irregularities in the lip and neck of the container. However, it must not fit so snugly that it is difficult to introduce into the neck of the container. Closures may be inserted aseptically with sterile forceps or directly with hands encased in sterile rubber gloves. When rubber closures are to be inserted mechanically, their surface is often halogenated or treated with silicone to make them easier to insert. Thus, it is possible to convey the closure through a shute to the place where it is positioned over a vial and then inserted by a plunger or some other pressure device. An example of such a mechanical device is shown in Fig 84-23. Mechanical stoppering has been developed to meet the need for high-speed production.

Rubber closures are held in place by means of aluminum caps. The caps cover the closure and are crimped under the lip of the vial or bottle to hold them in place (see Fig 84-4). The closure cannot be removed without destroying the aluminum cap. Therefore, an intact aluminum cap is proof that the closure has not been removed intentionally or unintentionally. Such confirmation is necessary to assure the integrity of the contents as to sterility and other aspects of quality.

The aluminum caps are so designed that the outer layer of

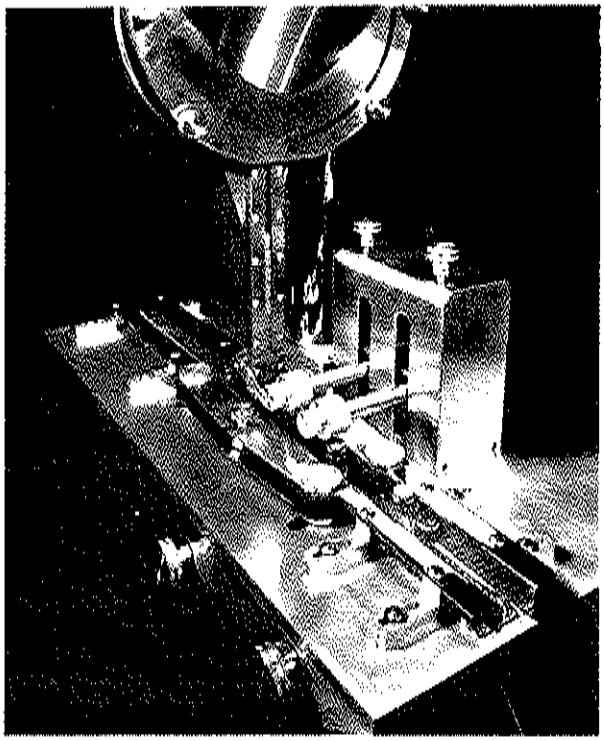


Fig 84-23. Mechanical device for inserting rubber closures in vials (courtesy, Perry).

double-layered caps, or the center of single-layered caps, can be removed to expose the center of the rubber closure without disturbing the band which holds the closure in the container. Rubber closures for use with intravenous administration sets often have a permanent hole through the closure. In such cases, a thin rubber disk overlayed with a solid aluminum disk is placed between an inner and outer aluminum cap, thereby providing a seal of the hole through the closure. These are called triple-layered aluminum caps.

Single-layered aluminum caps may be applied by means of a hand crimper known as the Fermpress (Suppliers: *West, Wheaton*). Double- or triple-layered caps require greater force for crimping; therefore, heavy-duty mechanical crimpers (see Fig 84-24) are required (Suppliers: *Cozzoli, Perry, Seidenader, West, Wheaton*).

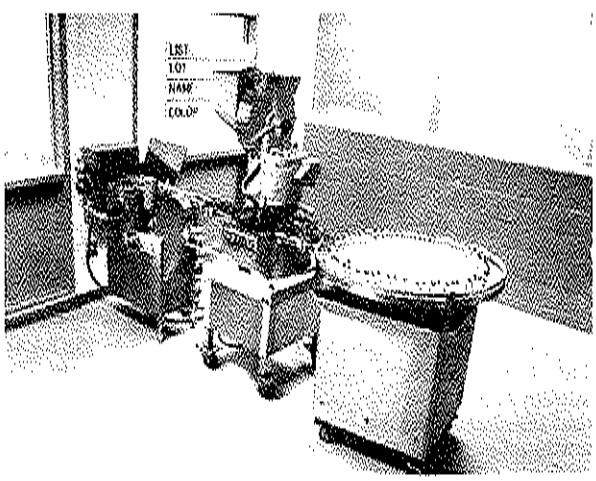


Fig 84-24. Applying aluminum caps to vials at the end of the process line (courtesy, Abbott).

Sterilization

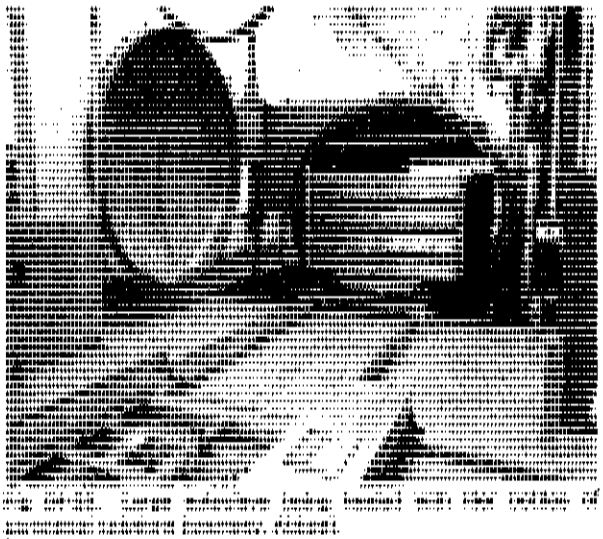
Whenever possible, the parenteral product should be sterilized after being sealed in its final container (terminal sterilization) and within as short a time as possible after the filling and sealing have been completed. Since this usually involves a thermal process, due consideration must be given to the effect of the elevated temperature upon the stability of the product. Many products, both pharmaceutical and biological, will be affected adversely by the elevated temperatures required for thermal sterilization. Such products must, therefore, be sterilized by a nonthermal method. Most thermolabile solutions may be sterilized by filtration through bacteria-retaining filters. Subsequently, all operations must be carried out in an aseptic manner so that contamination will not be introduced into the filtrate. To perform such an aseptic procedure is difficult, and the degree of its accomplishment is always uncertain. Colloids, oleaginous solutions, suspensions and emulsions that are thermolabile may require a process in which each component is sterilized and the product is formulated and processed under aseptic conditions. Because of the ever-present risk of a momentary or prolonged lapse in aseptic control during an aseptic process, and the dangerous condition that could result, sterilization of a product in its final container is preferred, if possible.

Some of the newer nonthermal methods of sterilization are finding important application to components of injections and administration devices. Certain dry solids such as penicillin, streptomycin, polyvitamins and certain hormones are being sterilized effectively by ionized radiations without adverse effects. Catgut sutures now are being sterilized routinely in the final package by this method. Administration sets, disposable needles and syringes and other plastic and stainless-steel equipment and components are being sterilized by ionizing radiations and by gaseous ethylene oxide sterilization. Generally speaking, however, neither of these methods may be used for liquid preparations without adverse effects on the product, and gaseous sterilization cannot be used where a glass container or other impervious barrier prevents the gas from permeating the material.

Dry-heat sterilization may be employed for a few dry solids that are not affected adversely by the high temperatures and for the relatively long heating period required. This method is applied most effectively to the sterilization of glassware and metalware. After sterilization, the equipment will be sterile, dry and, if the sterilization period is long enough, pyrogen-free.

Saturated steam under pressure (autoclaving) is the most commonly used and the most effective method for the sterilization of aqueous liquids or substances that can be reached or penetrated by steam.

Figure 84-25 shows liter containers of solution being loaded into an autoclave for sterilization. It is ineffective in anhydrous conditions, such as within a sealed ampul containing a dry solid or an anhydrous oil. Since the temperature employed in an autoclave is lower than that for dry-heat sterilization, equipment made of materials such as rubber and polypropylene may be sterilized if the time and temperature are controlled carefully. As mentioned previously, some injections will be affected adversely by the elevated temperature required for autoclaving. For some products, such as Dextrose Injection, the use of an autoclave designed to permit a rapid rise to sterilizing temperature and rapid cooling with water spray after the sterilizing hold-period will make it possible to use this method. Other products that will not withstand autoclaving temperatures may withstand marginal thermal methods such as tyndallization or inspissation. These methods may be rendered more effective for



some injections by the inclusion of a bacteriostatic agent in the product.

It should be obvious that all materials subjected to sterilization must be protected from subsequent contamination to maintain their sterile state. Therefore, they must be wrapped or covered so that microorganisms may not gain access when removed from the autoclave. Equipment and supplies are wrapped most frequently with paper and tied or sealed with special autoclave tape. The wrapping must permit penetration of steam during autoclaving but screen out microorganisms when dry. A double wrapping with lint-free parchment paper designed for such use is probably best. Synthetic fiber cloth such as nylon or Dacron also may be used for the inner wrapping. The openings of equipment subjected to dry-heat sterilization are often covered with silver–aluminum foil or with metal or glass covers. Cellulose wrapping materials are affected adversely by the high temperatures of dry-heat sterilization.

The effectiveness of any sterilization technique must be proved (validated) before it is employed; controls then being established to show that subsequent processes repeat the conditions proven to be effective. Since the goal of sterilization is to kill microorganisms, the ideal indicator to prove the effectiveness of the process is a biological one; resistant spores. However, many feel considerable hesitation about using biological indicators (BIs) during the processing of products because of the inherent risk of inadvertent contamination of the product or the environment. Also, it has been found that the resistance of spores may vary from lot to lot, thereby possibly giving false indications of reliability. However, today commercially prepared BIs are established as reliable for use in conjunction with physicial-parameter measurement for validating and monitoring sterilization processes. Such physical-parameter monitors include recording thermocouples, color-change indicators and melting indicators. This type of confirmatory evidence is an essential part of the sterilization record for a product.

Further details concerning methods of sterilization and their application will be found in Chapter 78. In addition, the USP provides suggestions concerning the sterilization of injections and related materials.

Freeze-Drying

Freeze-drying (lyophilization) is a process of drying in which water is sublimed from the product after it is frozen.[18] The particular advantages of this process are that biologicals and pharmaceuticals which are relatively unstable in aqueous solution can be processed and filled into dosage containers in the liquid state, taking advantage of the relative ease of processing a liquid. They can be dried without elevated temperatures, thereby eliminating adverse thermal effects, and stored in the dry state in which there are relatively few stability problems.

Further advantages are that these products are often more soluble and/or more rapidly soluble, dispersions are stabilized throughout their shelf life and products subject to degradation by oxidation have enhanced stability because the process is carried out in a vacuum.

However, the increased time and handling required for processing and the cost of the equipment limit the use of this process to those products which significantly have enhanced stability if stored in the dry state.

The fact that ice will sublime at pressures below 3 torr has been a long-established laboratory principle. The extensive program for freeze-drying human plasma during World War II provided the impetus for the rapid development of the process.

Freeze-drying essentially consists of the following:

1. Freezing an aqueous product at a temperature below its eutectic temperature.
2. Evacuating the chamber, usually below 0.1 torr (100 μm Hg).
3. Subliming ice on a cold condensing surface at a temperature below that of the product, the condensing surface being within the chamber or in a connecting chamber.
4. Introducing heat to the product under controlled conditions, thereby providing energy for sublimation at a rate designed to keep the product temperature below its eutectic temperature.

Figure 84-26 shows such a system. The product may be frozen on the shelf in the chamber by circulating refrigerant (usually Freon, ammonia or ethylene glycol) from the compressor through pipes within the shelf. After freezing is complete, which may require several hours, the chamber and condenser are evacuated by the vacuum pump, the condenser surface having been chilled previously by circulating refrigerant from the large compressor.

Heat then is introduced from the shelf to the product by electric resistance coils or by circulating hot water, silicone or glycol. The process continues until the product is dry (usually 1% or less moisture), leaving a sponge-like matrix of

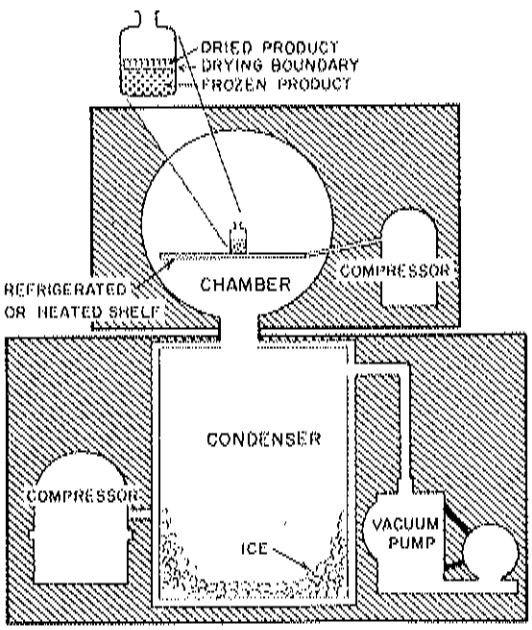


Fig 84-26. Essential components of a freeze-drying system.

the solids originally present in the product, the input of heat being controlled so as not to degrade the product.

For most pharmaceuticals and biologicals the liquid product is sterilized by filtration and then filled into the dosage container aseptically. The containers must remain open during the drying process; therefore, they must be protected from contamination during transfer from the filling area to the freeze-drying chamber, while in the freeze-drying chamber and at the end of the drying process until sealed.

The chambers may be equipped with hydraulic or rubber diaphragm internal-stoppering devices designed to push slotted rubber closures into the vials to be sealed while the chamber is still evacuated, the closures having been partially inserted immediately after filling so that the slots were open to the outside.

If internal stoppering is not available or containers such as ampuls are used, filtered dry air or nitrogen must be introduced to the chamber at the end of the process to establish atmospheric pressure. Then the containers must be removed and sealed under aseptic conditions. If the product is very sensitive to moisture, the environmental humidity also must be controlled until it is sealed.

Factors Affecting the Process Rate—The greater the depth of the product in the container, the longer will be the drying process. Therefore, a product to be frozen by placing the container on a refrigerated shelf (plug freezing) should be filled to a planned, limited depth. If a large volume of solution must be processed, the surface area may be increased and the depth decreased by freezing the solution on a slant or while rotating the container on an angle (shell freezing) in a liquid refrigerant bath, such as dry ice and alcohol.

The actual driving force for the process is the vapor pressure differential between the vapor at the surface where drying of the product is occurring (the drying boundary) and that at the surface of the ice on the condenser. The latter is determined by the temperature of the condenser as modified by the insulating effect of the accumulated ice. The former is determined by a number of factors, including:

1. The rate of heat conduction through the container and the frozen material, both usually relatively poor thermal conductors, to the drying boundary while maintaining all of the product below its eutectic temperature.
2. The impeding effect of the increasing depth of dried, porous product above the drying boundary.
3. The temperature and heat capacity of the shelf itself.

This may be visualized by referring to Fig 84-26.

The passageways between the product surface and the condenser surface must be wide open and direct for effective operation. Therefore, the condensing surfaces in large freeze-driers are usually in the same chamber as the product. Evacuation of the system is necessary to reduce the impeding effect that collisions with air molecules would have on the passage of water molecules. However, the residual pressure in the system must be greater than the vapor pressure of the ice on the condenser or the ice will be vaporized and pulled into the pump, an event detrimental to most pumps.

The amount of solids in the product, their particle size and their thermal conductance will affect the rate of drying. The more solids present, the more impediment will be provided to the escape of the water vapor. The smaller the particle size, particularly the crystal size of the ice, the faster the drying generally will be. The poorer the thermal conducting properties of the solids in the product, the slower will be the rate of heat transfer through the frozen material to the drying boundary.

The rate of drying is essentially slow, most often requiring 24 hr or longer for completion. The actual time required, the rate of heat input and the product temperatures that may be used must be determined for each product and then reproduced carefully with successive processes.

Factors Affecting Formulation—The active constituent of many pharmaceutical products is present in such a small quantity that if freeze-dried alone its presence would be hard to detect visually. Therefore, excipients are often added to increase the amount of solids.

Some consider it ideal for the dried-product plug to occupy essentially the same volume as that of the original solution. To achieve this, the solids content of the original product must be between approximately 5 and 25%. Among the substances found most useful for this purpose, usually as a combination, are sodium or potassium phosphates, citric acid, tartaric acid, gelatin and carbohydrates such as dextrose, mannitol and dextran.

Each of these substances contributes appearance characteristics to the plug, such as whether dull and spongy or sparkling and crystalline, firm or friable, expanded or shrunken and uniform or striated. Therefore, the formulation of a product to be freeze-dried must include consideration not only of the nature and stability characteristics required during the liquid state, both freshly prepared and when reconstituted before use, but the characteristics desired in the dried plug.

Modifications in the Process and Equipment—In some instances a product may be frozen in a bulk container or in trays rather than in the final container and then handled as a dry solid. This may be desirable when large volumes of a product are processed.

Heat may be introduced to all sides of the product by radiation from infrared sources, rather than only from the bottom as with conductive heating. While this generally increases the rate of drying, there are at least two major disadvantages to radiant heating of pharmaceuticals; these are (1) multiple containers produce shadowing with resultant blockage of the radiations and (2) the dried material on the outside of the frozen product may be scorched easily by the heat as drying progresses.

When large quantities of material are processed it may be desirable to use ejection pumps in the equipment system. These draw the vapor into the pump and eject it to the outside, thereby eliminating the need for a condensing surface. Such pumps are expensive and usually practical only in large installations.

Available freeze-driers (Suppliers: *Edwards, FTS, Hull, NRC, Stokes, Virtis*) range in size from small laboratory units to large industrial models such as the one shown in Fig 84-27. Their selection requires consideration of such factors as:

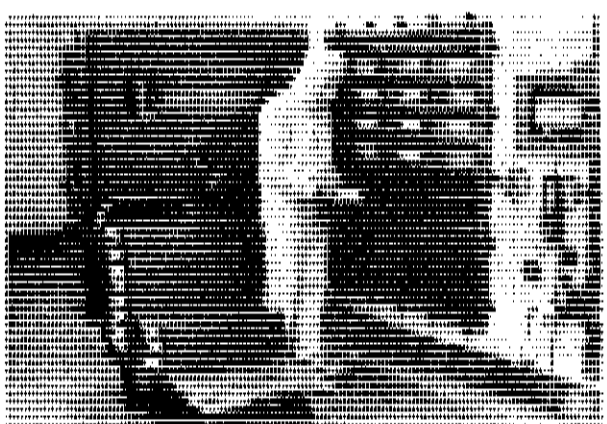


Fig 84-27. Aseptic loading of freeze-drier (courtesy, Upjohn).

The tray area required.
The volume of water to be removed.
Whether or not aseptic processing will be involved.
Is internal stoppering required?
Will separate freezers be used for initial freezing of the product.
The degree of automatic operation desired.

Other factors involved in the selection and use of equipment are considered in the literature.[19]

Freeze-drying is now being used for research in the preservation of human tissue and is finding increasing application in the food industry. Progress on new developments is being made in both the process and the equipment.[20]

Quality Assurance and Control

The importance of undertaking every possible means to assure the quality of the finished product cannot be overemphasized. Every component and step of the manufacturing process must be subjected to intense scrutiny to be confident that quality is attained in the finished product. The responsibility for achieving this quality is divided appropriately in concept and practice into Quality Assurance (QA) and Quality Control (QC). QA relates to the studies made and the plans developed for assuring quality of a product prospectively. QC embodies the carrying out of these plans and includes all of the tests and evaluations performed to be sure that quality has been achieved in a specific lot of product.

The principles for achieving quality are basically the same for the manufacture of any pharmaceutical. These are discussed in Chapter 82. During the discussion of the preparation of injections, mention was made of numerous quality requirements for components and manufacturing processes. Here, only certain tests characteristically applicable to the finished parenteral products will be discussed.

Sterility Test

All lots of injections in their final containers must be tested for sterility. The USP prescribes the requirements for this test for official injections. The FDA uses these requirements as a guide for testing unofficial sterile products. The official test has acknowledged limitations in the information that it can provide. Therefore, it should be noted that this test is not intended as a thoroughly evaluative test for a product subjected to a sterilization method of unknown effectiveness. It is intended primarily as a check test on the probability that a previously validated sterilization procedure has been repeated, or to give assurance of its continued effectiveness. A discussion of sterility testing is given in Chapter 78.

It should be noted that a "lot" with respect to sterility testing is that group of product containers which has been subjected to the same sterilization procedure. For containers of a product which have been sterilized by autoclaving, for example, a lot would constitute those processed in a particular sterilizer cycle. For an aseptic filling operation, a lot would constitute all of those product containers filled during a period when there was no change in the filling assembly or equipment and which is no longer than one working day or shift.

Pyrogen Test

The presence of pyrogens in parenteral preparations is evaluated by a qualitative fever response test in rabbits. The USP tests are described in Chapter 27. Rabbits are used as test animals because they show a physiological response to pyrogenic substances similar to that by man. While a minimum pyrogenic dose (MPD), the amount just sufficient to cause a positive USP Pyrogen Test response, sometimes may produce uncertain test results, a content equal to a few times the MPD will leave no uncertainty. Therefore, the test is valid and has continued in use since introduced by Seibert in 1923. It should be understood that not all injections may be subjected to the rabbit test since the medicinal agent may have a physiological effect on the test animal such that any fever response would be masked.

A new test for pyrogens recently has been accepted, not only for in-process control for pharmaceutical products but also for release testing of such products and for devices. It is an *in vitro* test based on the gelling or color development of a pyrogenic preparation in the presence of the lysate of the amebocytes of the horseshoe crab (*Limulus polyphemus*). The Limulus Test, as it is called, is simpler, more rapid and of greater sensitivity than the rabbit test.[21] Although it detects only the endotoxic pyrogens of Gram-negative bacteria, this probably will not limit its use significantly since most environmental contaminants gaining entrance to sterile products are Gram-negative. The test has gained in stature to the point that automated techniques have been developed.[22]

Particulate Evaluation

Particulate matter in parenteral solutions long has been recognized as unacceptable since the user could be expected to conclude that the presence of visible "dirt" would suggest that the product is of inferior quality. Today, it is recognized that the presence of particles in solution, particularly if injected intravenously, can be harmful. While data defining the extent of risk and the effects produced still are limited, it has been shown that particles of lint, rubber, insoluble chemicals and other foreign matter can produce emboli in the vital organs of animals and man.[23] Further, it has been shown that the development of infusion-phlebitis may be related to the presence of particulate matter in intravenous fluids.[24]

The particle size of particular concern has not been clearly delineated, but it has been suggested that since erythrocytes have a diameter of approximately 4.5 μm, particles of more than 5 μm should be the basis for evaluation. This is a considerably smaller particle than can be seen with the unaided eye; approximately 50 μm is the lower limit unless the Tyndall effect is used whereby particles as small as 10 μm can be seen by the light scattered from them.

The USP specifies that good manufacturing practice requires that each final container of an injection be subjected individually to a visual inspection and that containers in which visible particles can be seen should be discarded. This 100% inspection of a lot of product is designed to prevent the distribution and use of parenterals which contain particulate matter that may be harmful psychologically or organically to the participant. Therefore, all of the product units from a production line currently are being inspected individually by human inspectors under a good light, baffled against reflection into the eye and against a black-and-white background. This inspection is subject to the limitation of the size of particles that can be seen, the variation of visual acuity from inspector to inspector, their emotional state, eye strain, fatigue and other personal factors that will affect what is seen. However, it does provide a means for eliminating the few units which normally contain visible particles.

Since it is recognized that visual inspection will not detect the presence of particles smaller than approximately 50 μm in size, the USP has established a microscopic test method

for identifying particles in large-volume intravenous solutions and has set limits of not more than 50 particles/mL of 10 μm and larger in size and not more than 5 particles/mL of 25 μm and larger in size. This method consists essentially of filtering a measured sample of solution through a membrane filter under ultraclean conditions and then counting the particles on the surface of the filter using oblique light, under a microscope, at both 40× and 100× magnification. These standards are being met readily by the large-volume parenteral solutions currently being manufactured in the US.

More recently the USP established standards for small-volume parenterals to be given intravenously, using an electronic instrument that counts and measures the size of particles by means of a shadow cast by the particle as it passes through a high-intensity light beam (Suppliers: *Climet*, *HIAC*). The limits prescribed are not more than 10,000 particles/mL of ≥ 10 μm in size and not more than 1000 particles/mL ≥ 25 μm in size. These specifications were developed on the premise that as many as five such products may be added to a 1-L bottle of a large-volume parenteral and five products should not contribute more than the overall limits of particles prescribed for a large-volume parenteral. Whether or not these standards are realistic toxicologically has not been established; rather, the objective of the compendium is to establish specification limits that would encourage the preparation of clean parenteral solutions, particularly for those to be given intravenously.

It also should be realized that administration sets and the techniques used in the hospital for preparing and administering intravenous infusion fluid may introduce substantial amounts of particulate matter into an otherwise clean solution. Therefore, the pharmaceutical manufacturer, the administration set manufacturer, the hospital pharmacist, the nurse and the physician must share responsibilities for making sure that the patient receives a clean intravenous injection.

The USP methods for counting and sizing particulate matter in intravenous solutions are not the only methods available for such determinations. A number of electronic particle counters are available that use the light-scattering principle to count particles in a liquid sample (Suppliers: *Climet, Met One, Royco*). There also is an instrument available which counts particles and sizes them by measuring the effect on the resistance between two electrodes as the particles pass between them (Supplier: *Coulter*). It is obvious that only the visual inspection can be used for in-line evaluation of every container produced commercially. All of these methods require very stringent ultraclean preparation techniques to assure reasonable accuracy in counting and sizing only the particles in the solution, rather than those that may have been introduced inadvertently during the sample preparation or the testing procedure. Further, these test procedures are destructive and, therefore, can be performed only on samples of the production lot. Further information may be found in a review article.[25]

Leaker Test

Ampuls that have been sealed by fusion must be subjected to a test to determine whether or not a passageway remains to the outside; if so, all or a part of the contents may leak to the outside and spoil the package, or microorganisms or other contaminants may enter. Changes in temperature during storage cause expansion and contraction of the ampul and contents, and will accentuate interchange if a passageway exists, even if microscopic in size.

This test usually is performed by producing a negative pressure within an incompletely sealed ampul while the ampul is entirely submerged in a deeply colored dye solution. Most often, approximately a 1% methylene blue solution is employed. The test may be performed by subjecting the ampuls to a vacuum in a vacuum chamber, the ampuls being submerged in a dye bath throughout the process. Another procedure frequently employed is to simply autoclave the ampuls in a dye bath. A modification of this is to remove them from the autoclave while hot and quickly submerge them in a cool bath of dye solution. After carefully rinsing the dye solution from the outside, color from the dye will be visible within a leaker. Leakers, of course, are discarded.

Vials and bottles are not subjected to a leaker test because the sealing material (rubber stopper) is not rigid. Therefore, results from such a test would be meaningless. However, evacuated bottles containing a liquid may be checked for a sharp "click" sound produced when struck with an implement such as a rubber mallet.

Safety Test

The National Institutes of Health requires of most biological products routine safety testing in animals. Under the Kefauver-Harris Amendments to the Federal Food, Drug, and Cosmetic Act, most pharmaceutical preparations are now required to be tested for safety. Because it is entirely possible for a parenteral product to pass the routine sterility test, pyrogen test and chemical analyses and still cause unfavorable reactions when injected, a safety test in animals is essential to provide additional assurance that the product does not have unexpected toxic properties. Safety tests in animals are discussed in detail in the USP.

Packaging and Labeling

A full discussion of the packaging of parenteral preparations is beyond the scope of this text. It is essential, of course, that the packaging should provide ample protection for the product against physical damage from shipping, handling and storage as well as protecting light-sensitive materials from ultraviolet radiation. An extensive review of this subject has been published.[26]

Packaging—The USP includes certain requirements for the packaging and storage of injections, as follows:

1. The volume of injection in single-dose containers is defined as that which is specified for parenteral administration at one time and is limited to a volume of 1 L.
2. Parenterals intended for intraspinal, intracisternal or peridural administration are packaged only in single-dose containers.
3. Unless an individual monograph specifies otherwise, no multiple-dose container shall contain a volume of injection more than sufficient to permit the withdrawal and administration of 30 mL.
4. Injections packaged for use as irrigation solutions or for hemofiltration or dialysis are exempt from the foregoing requirements relating to packaging. Containers for injections packaged for use as hemofiltration or irrigation solutions may be designed to empty rapidly and may contain a volume in excess of 1 L.
5. Injections intended for veterinary use are exempt from the packaging and storage requirements concerning the limitation to single-dose containers and to volume of multiple-dose containers.

Labeling—The labeling of an injection must provide the physician or other user with all of the information needed to assure the safe and proper use of the therapeutic agent. Since all of this information cannot be placed on the immediate container and be legible, it may be provided on accompanying printed matter. General labeling requirements for drugs are discussed in Chapter 107.

A restatement of the labeling definitions and requirements of the USP for Injections is as follows:

'The term "labeling" designates all labels and other written, printed or graphic matter upon an immediate container or upon, or in, any package or wrapper in which it is enclosed, with the exception of the outer shipping container. The term "label" designates that part of the labeling upon the immediate container.

The label states the name of the preparation, the percentage content of drug of a liquid preparation, the amount of active ingredient of a dry preparation, the volume of liquid to be added to prepare an injection or suspension from a dry preparation, the route of administration, a statement of storage conditions and an expiration date. Also, the label must indicate the name of the manufacturer or distributor and carry an identifying lot number. The lot number is capable of providing access to the complete manufacturing history of the specific package, including each single manufacturing step.

The container label is so arranged that a sufficient area of the container remains uncovered for its full length or circumference to permit inspection of the contents.

The label must state the name of the vehicle and the proportions of each constituent, if it is a mixture; the names and proportions of all substances added to increase stability or usefulness and the expiration date where required by the individual monograph.

Preparations labeled for use as dialysis, hemofiltration or irrigation solutions must meet the requirements for Injections other than those relating to volume and also must bear on the label statements that they are not intended for intravenous injection.

Injections intended for veterinary use are so labeled.

References

1. Griffenhagen GB: *Bull Parenter Drug Assoc 16(2):* 12, 1962.
2. Tousignant DR: *Am J Hosp Pharm 34:* 943, 1977.
3. *Accreditations Manual for Hospitals* JCAH, Chicago, 1987.
4. Bergman DH: *Drug Intel Clin Pharm 11:* 345, 1977.
5. 21 CFR 210, Current Good Manufacturing Practice in manufacturing, processing, packing or holding of drugs; general, Supt. of Documents, US GPO, Washington DC, 1988.
6. Spiegel AJ, Noseworthy MM: *J Pharm Sci 52:* 917, 1963.
7. Schroeter LC: *Ibid 50:* 891, 1961.
8. Morton DK: *J Parenter Sci Technol 41:* 145, 1987.
9. Autian J: *Bull Parenter Drug Assoc 22:* 276, 1968.
10. *Tech Methods Bull No 3,* PDA, Philadelphia, 1982.
11. *Tech Methods Bull No 2,* PDA, Philadelphia, 1981.
12. Federal Std No 209C, GSA, Washington DC 20407, March 1988.
13. Whyte W, Niven L: *J Parenter Sci Technol 40:* 182, 1986.
14. Anschel J: *Bull Parenter Drug Assoc 31:* 47, 1977.
15. Grimes TL, Fonner DE *et al: Ibid 31:* 179, 1977.
16. Myers T, Chrai S: *J Parenter Sci Technol 35:* 8, 1981.
17. Singh M, Ruvin LJ: *Ibid 40:* 34, 1986.
18. *Ann NY Acad Sci 85:* 501–734, 1965.
19. Morgan SL, Spotts MR: *Pharm Technol 3:* 94–101, 114, 1979.
20. Nail SL, Gatlin LA: *J Parenter Sci Technol 39:* 16, 1985.
21. Cooper JF: *Bull Parenter Drug Assoc 29:* 122, 1975.
22. Novitsky TJ, Ryther SS *et al: J Parenter Sci Technol 36:* 11, 1982.
23. Garvan JM, Gunner BW: *Med J Aust 2:* 1, July 4, 1964.
24. Deluca P *et al: Am J Hosp Pharm 32:* 1001, 1975.
25. Borchert SJ, Abe A *et al: J Parenter Sci Technol 40:* 212, 1986.
26. *Tech Report No 5,* PDA, Philadelphia, 1984.

Bibliography

Akers MJ: *Parenteral Quality Control,* Marcel Dekker, New York, 1985.

Avis KE, Lavchuk JW, In King RE, ed: *Dispensing of Medication,* 9th ed, Mack Publ Co, Easton PA, Chap 9, 1984.

Avis KE, Lachman L, Lieberman HA, eds: *Pharmaceutical Dosage Forms: Parenteral Medications,* vol 1, Marcel Dekker, New York, 1984.

Ibid, vol 2, Marcel Dekker, New York, 1986.

Avis KE. In Lachman L *et al: The Theory and Practice of Industrial Pharmacy,* 3rd ed, Lea & Febiger, Philadelphia, Chaps 21 and 22, 1986.

Block SS, ed: *Disinfection, Sterilization and Preservation,* 3rd ed, Lea & Febiger, Philadelphia, 1983.

Carleton FJ, Agalloco JP, eds: *Validation of Aseptic Pharmaceutical Processes,* Marcel Dekker, New York, 1986.

Gaughran ERL, Kereluk K, eds: *Sterilization of Medical Products,* Johnson & Johnson, New Brunswick NJ, 1977.

Martin EW *et al: Techniques of Medication,* Lippincott, Philadelphia, 1969.

Meltzer TH, ed: *Filtration in the Pharmaceutical Industry,* Marcel Dekker, New York, 1987.

Meryman HT, ed: *Cryobiology,* Academic, New York, 1966.

Pearson FC III: *Pyrogens,* Marcel Dekker, New York, 1985.

Phillips GB, Miller WS, eds: *Industrial Sterilization,* Duke Univ Press, Durham NC, 1973.

Turco S, King RE: *Sterile Dosage Forms,* 3rd ed, Lea & Febiger, Philadelphia, 1987.

CHAPTER 85

Intravenous Admixtures

Salvatore J Turco, PharmD

Professor of Pharmacy
Temple University School of Pharmacy
Philadelphia, PA 19140

It has been estimated that 40% of all drugs administered in hospitals are given in the form of injections and their use is increasing. Part of this increase in parenteral therapy is due to the wider use of intravenous fluids (IV fluids). In the last decade the use of IV fluids has doubled, increasing from 150 million units to 300 million units annually. Not only do IV fluids continue to serve as the means for fluid replacement, electrolyte-balance restoration and supplementary nutrition, but they also are playing major roles as vehicles for administration of other drug substances and in total parenteral nutrition (TPN). TPN fluids are finding greater use as the means of administering other drugs because of convenience, the means of reducing the irritation potential of the drugs and the desirability for continuous and intermittent drug therapy. The techniques for providing TPN parenterally have improved steadily in the last decade, and such use is increasing markedly. The use of IV fluids for these purposes requires the compounding of specific intravenous admixtures (parenteral prescriptions) to meet the clinical needs of a given patient. However, the combination of drug substances in an IV fluid can promote parenteral incompatibilities and give rise to conditions not favorable for drug stability. A new area of specialization has been created for hospital pharmacists who can develop the expertise to prepare these solutions—recognizing their compatibility and stability problems and the potential for contamination—and participate in the administration of the solutions. The complex compounding of an order for TPN requires knowledgeable personnel capable of making accurate calculations, compounding and having aseptic technique. The parenteral prescription is becoming increasingly important in hospitals. Centralized admixture programs are now found in 70% of the nation's hospitals having 300 beds or more. Equipment available for administering IV fluids has become more sophisticated, and has made possible increased accuracy of dosage and led to the development of new concepts and methods of nutrition and drug therapy.

Intravenous Fluids

Large-volume injections intended to be administered by intravenous infusion commonly are called IV fluids and are included in the group of sterile products referred to as large-volume parenterals. These consist of single-dose injections having a volume of 100 mL or more and containing no added substances. Intravenous fluids are packaged in containers having a capacity of 100 to 1000 mL. Minitype infusion containers of 250-mL capacity are available with 50- and 100-mL partial fills for solution of drugs when used in the "piggyback" technique (ie, the administration of a second solution through a Y-tube or gum-rubber connection in the administration set of the first intravenous fluid, thus avoiding the need for another injection site). In addition to the IV fluids, this group also includes irrigation solutions and solutions for dialysis.

Intravenous fluids are sterile solutions of simple chemicals such as sugars, amino acids or electrolytes—materials which easily can be carried by the circulatory system and assimilated. Prepared with Water for Injection USP, the solutions are pyrogen-free. Because of the large volumes administered intravenously, the absence of particulate matter assumes a significant role in view of possible biological hazards resulting from insoluble particles. Absence of particulate matter or clarity of IV fluids is as important at the time of administration following their manipulation in the hospital as it is at the time of manufacture of the injection.

Limits for particulate matter occurring in IV fluids, or large-volume injections used for single-dose infusion, are defined in the USP. This represents the first regulatory attempt to define limits for particulate matter in parenterals. Limits also apply to multiple-dose injections, small-volume injections or injections prepared by reconstitution from sterile solids. The USP defines particulate matter as extraneous, mobile, undissolved substances, other than gas bubbles, unintentionally present in parenteral solutions. The total numbers of particles having effective linear dimensions equal to or larger than 10 μm and larger than 25 μm are counted. The IV fluid meets the requirement of the test if it contains not more than 50 particles per mL which are equal to or larger than 10 μm, and not more than 5 particles per mL which are equal to or larger than 25 μm in linear dimension.

Intravenous fluids commonly are used for a number of clinical conditions. These include

Correction of disturbances in electrolyte balance.
Correction of disturbances in body fluids (fluid replacement).
The means of providing basic nutrition.
The basis for the practice of providing TPN.
Use as vehicles for other drug substances.

In both of the latter two cases it has become common practice to add other drugs to certain IV fluids to meet the clinical needs of the patient. Using IV fluids as vehicles offers the advantages of convenience, the means of reducing the irritation potential of the drug and a method for continuous drug therapy. However, the practice requires that careful consideration be given to the stability and compatibility of additives present in the IV fluids serving as the vehicles. This approach also demands strict adherence to aseptic techniques in adding the drugs, as well as in the administration of the IV fluids. These procedures are discussed later in the chapter. The IV fluids commonly used for parenterals are shown in Table I.

Many disease states result in electrolyte depletion and loss. Proper electrolyte concentration and balance in plasma and tissues are critical for proper body function. Electrolyte restoration and balance are achieved most rapidly through administration of IV fluids. Required electrolytes include sodium and chloride ions, which in normal saline more closely approximate the composition of the extracellular fluid than solutions of any other single salt; potassium, the principal intracellular cation of most body tissues and an essential for the functioning of the nervous and muscular

Table I—Fluids Used Commonly for IV Use

Injection	Concentration (%)	pH	Therapeutic Use
Alcohol			
with D5/W[a]	5	4.5	Sedative, analgesic, calories
with D5/W in NSS[b]	5		Sedative, analgesic, calories
Amino Acid (Synthetic)			Fluid and nutrient replenisher
Aminosyn II (*Abbott*)	3.5; 7	5.25	
FreAmine III (*McGaw*)	8.5	6.6	
Travasol (*Baxter*)	3.5; 5.5; 8.5	6.0	
Ammonium Chloride	2.14	4.5–6.0	Metabolic alkaloids
Dextran 40			Priming fluid for extracorporeal circulation
in NSS	10	5	
in D5/W	10	4	Priming fluid for extracorporeal circulation
Dextran 70			
in NSS	6	5	Plasma volume expander
in D5/W	6	4	Plasma volume expander
Dextrose (Glucose, D5/W)	2.5–50	3.5–6.5	Fluid and nutrient replenisher
Dextrose and Sodium Chloride	Varying concn of dextrose from 5–20 with varying concn of sodium chloride from 0.22–0.9	3.5–6.5	Fluid, nutrient and electrolyte replenisher
Invert Sugar (Fructose and Dextrose)	5, 10	4.0	Fluid and nutrient replenisher
Lactated Ringer's (Hartmann's)		6.0–7.5	Systemic alkalizer; fluid and electrolyte replenisher
NaCl	0.6		
KCl	0.03		
$CaCl_2$	0.02		
Lactate	0.3		
Mannitol	5	5.0–7.0	Osmotic diuresis
also in combination with dextrose or sodium chloride	10 15 20		
Multiple electrolyte solutions varying combinations of electrolytes, dextrose, fructose, invert sugar		5.5	Fluid and electrolyte replacement
Ringer's		5.0–7.5	Fluid and electrolyte replenisher
NaCl	0.86		
KCl	0.03		
$CaCl_2$	0.033		
Sodium Bicarbonate	5	8	Metabolic acidosis
Sodium Chloride	0.45; 0.9; 3; 5	4.5–7.0	Fluid and electrolyte replenisher
Sodium Lactate	1/6 *M*	6.3–7.3	Fluid and electrolyte replenisher
Sterile Water for Injection		5.5	Diluent

[a] 5% Dextrose in water.
[b] Normal Saline Solution.

systems as well as the heart; magnesium, as a nutritional supplement especially in TPN solutions and phosphate ion, important in a variety of biochemical reactions. In addition to the number of standard electrolyte fluids shown in Table I, a large number of combinations of electrolytes in varying concentrations are available commercially. Some of these electrolyte fluids also contain dextrose.

Dextrose Injection 5% (D5/W) is the most frequently used IV fluid, either for nutrition or fluid replacement. It is isotonic and administered intravenously into a peripheral vein; 1 g of dextrose provides 3.4 cal and 1 L of D5/W supplies 170 cal. The body utilizes dextrose at a rate of 0.5 g per kg of body weight per hr. More rapid administration can result in glycosuria. Therefore, 1 L of D5/W requires 1½ hours for assimilation. The pH range of D5/W can vary from 3.5 to 6.5. The wide range permitted is due to the free sugar acids present and formed during the sterilization and storage of the injection. To avoid incompatibilities when other drug substances are added to Dextrose Injection, the possible low pH should be considered in using it as a vehicle. More concentrated solutions of dextrose are available and provide increased calorie intake with less fluid volume. Being hypertonic, the more concentrated solutions may be irritating to peripheral veins. Highly concentrated solutions are administered in a larger central vein. Other IV fluids used for intravenous admixtures and providing calories include solutions containing invert sugar. There is some evidence that fructose, unlike dextrose, may be used in diabetic patients; the 10% injection is hypertonic and provides 375 cal per L. Invert sugar consists of equal parts of dextrose and fructose; it is claimed that the presence of fructose promotes more rapid utilization of dextrose.

Intravenous fluids containing crystalline amino acids can provide biologically usable amino acids for protein synthesis (Chapter 51). Protein contributes to tissue growth, wound repair and resistance to infection. The protein requirement for the normal adult is 1 g per kg per day; children and patients under stress require greater amounts. Attempts are made to maintain a positive nitrogen balance, indicating that the protein administered is being utilized properly and

not broken down and eliminated through the urine as creatinine and urea, which are normal waste products. In a positive nitrogen balance patients are taking in more nitrogen than they are eliminating. In a negative nitrogen balance there is more nitrogen being eliminated through the urine regularly than is being administered intravenously. This means that tissues are continuing to be torn down and repair is not necessarily taking place. Amino Acid Injection can afford the total body requirements for proteins by the procedure known as TPN (discussed below) or be used for supplemental nutrition by peripheral administration. In addition to the amino acids, these nutritional injections also may contain dextrose, electrolytes, vitamins and insulin. Fat emulsion (*Intralipid*, Kabi Vitrum AB; *Liposyn II*, Abbott and *Travamulsion*; Travenol) sometimes is used concurrently but usually administered at another site.

Packaging Systems

Containers for intravenous fluids must be designed to maintain solution sterility, clarity (freedom from particulate matter) and nonpyrogenicity from the time of preparation, through storage and during clinical administration. Container closures must be designed to facilitate insertion of administration sets through which the injections are administered, at a regulated flow-rate, into suitable veins. IV fluids are available in glass and plastic containers; the latter may be made from either a flexible or semirigid plastic material. IV fluids are supplied in 1000-mL, 500-mL and 250-mL sizes in addition to 250-mL capacity containers packaged with 50 or 100 mL of D5/W or Sodium Chloride Injection for piggyback use. IV fluids in glass containers are packaged under vacuum, which must be dissipated prior to use. For fluid to leave the IV glass container and flow through the administration set, some mechanism is necessary to permit air to enter the container. Current flexible plastic systems do not require air introduction in order to function. Atmospheric pressure pressing on the container forces the fluid to flow.

All glass and plastic containers are single-dose and should be discarded after opening even if not used. Intravenous fluids are packaged with approximately 3% excess fill to allow for removal of air from the administration set and permit the labeled volume to be delivered from the container. The containers are graduated at 20-mL increments on scales that permit the volume in container to be determined either from an upright or inverted position. Glass containers have aluminum and plastic bands for hanging, while plastic containers have eyelet openings or plastic straps for attachment to IV poles.

Table II—IV Fluid Systems

Source	Container	Characteristics
Baxter	Glass	Vacuum Air tube
Baxter (*Viaflex*)	Plastic	Polyvinyl chloride Flexible Nonvented
McGaw	Glass	Vacuum Air tube
McGaw (*Accumed*)	Plastic	Polyolefin Semirigid
Abbott	Glass	Vacuum Air filter[a]
Abbott (*Lifecare*)	Plastic	Polyvinyl chloride Flexible Nonvented

[a] Part of administration set.

Fluids for IV use are available from three sources; all provide both glass and plastic containers. The glass-container systems of Baxter and McGaw are similar. The characteristics of current packaging systems are summarized in Table II.

Administration Sets

Administration sets used to deliver fluids intravenously are sterile, pyrogen-free and disposable. Although these sets are supplied by different manufacturers, each for its own system, they have certain basic components. These include a plastic spike to pierce the rubber closure or plastic seal on the IV container, a drip (sight) chamber to trap air and permit adjustment of flow rate and a length (150 to 450 cm) of polyvinyl chloride tubing terminating in a gum-rubber injection port. At the tip of the port is a rigid needle or catheter adapter. An adjustable clamp (screw or roller type) on the tubing pinches the tubing to regulate flow. Since the gum-rubber port is self-sealing, additional medication can be added to the IV system at these ports of entry. Glass containers that have no air tubes require air-inlet filters designed as part of the administration set (Abbott). See Figs 85-1 to 85-5.

Administration Procedures

In the administration of IV fluids, the primary IV container provides for fluid replacement, electrolyte replenishment,

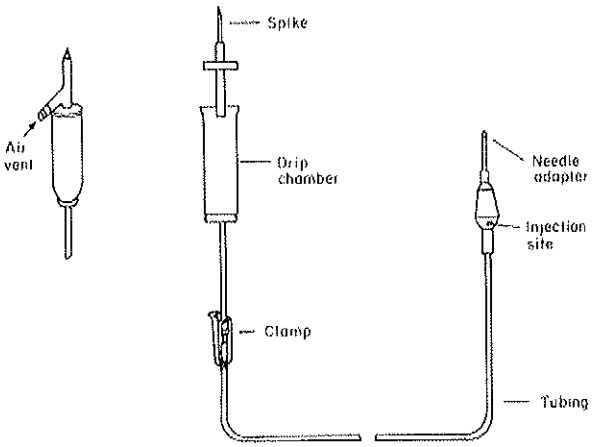


Fig 85-1. Parts of basic administration sets.

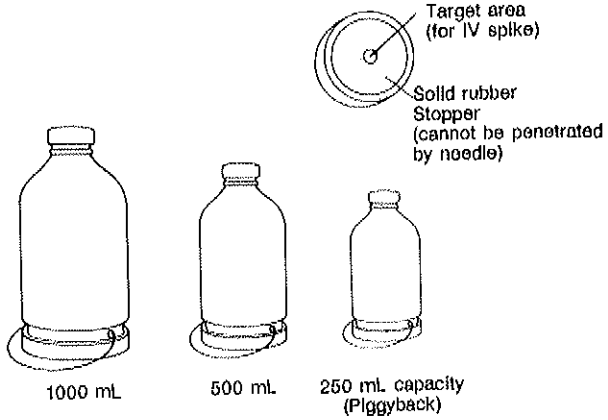


Fig 85-2. Abbott IV glass container. The air venting is provided through the air filter located in the spike of the administration set. See Fig 85-1.

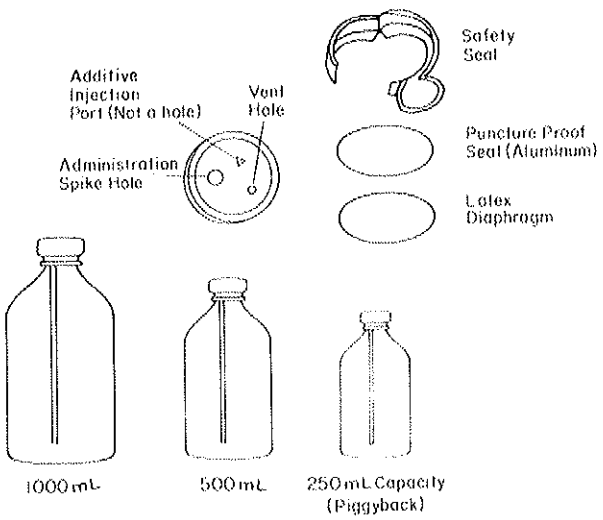


Fig 85-3. Baxter and McGaw glass containers. The plastic air tube allows the air to enter the bottle as the fluid is infused into the patient. The spike of the administration set is not vented. See Fig 85-1.

drug therapy or nutrition; the fluid can be infused over a 4- to 8-hr period. In some cases an IV fluid is infused slowly for the purpose of keeping the vein open (KVO). This will allow additional drugs to be administered when required. The primary IV fluid also can serve as a vehicle for other drugs to be administered, thus becoming an intravenous admixture (IV drip) and results in continuous blood levels of added drugs once the steady state has been reached.

In preparing an IV fluid for administration, the following procedure is used.

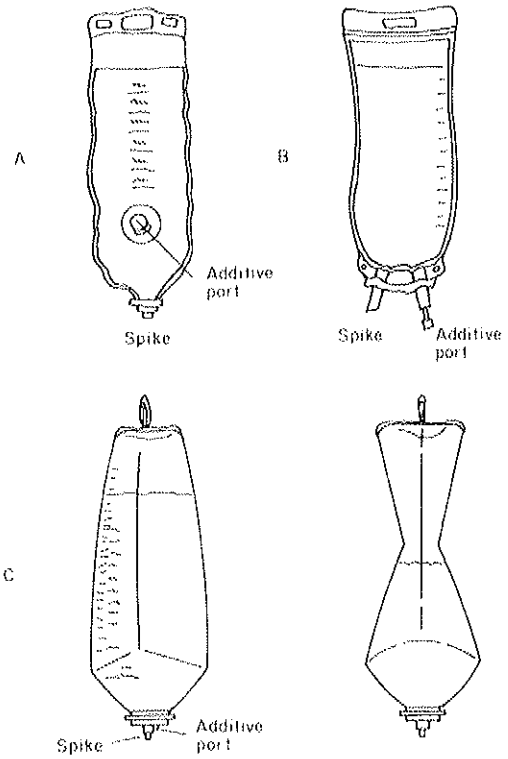


Fig 85-4. (A) Abbott (*Lifecare*) polyvinyl chloride flexible container; (B) Baxter (*Viaflex*) polyvinyl chloride flexible container; McGaw (*Accumed*) polyolefin semirigid container, front and side views. These containers take nonvented administration sets. See Fig 85-1.

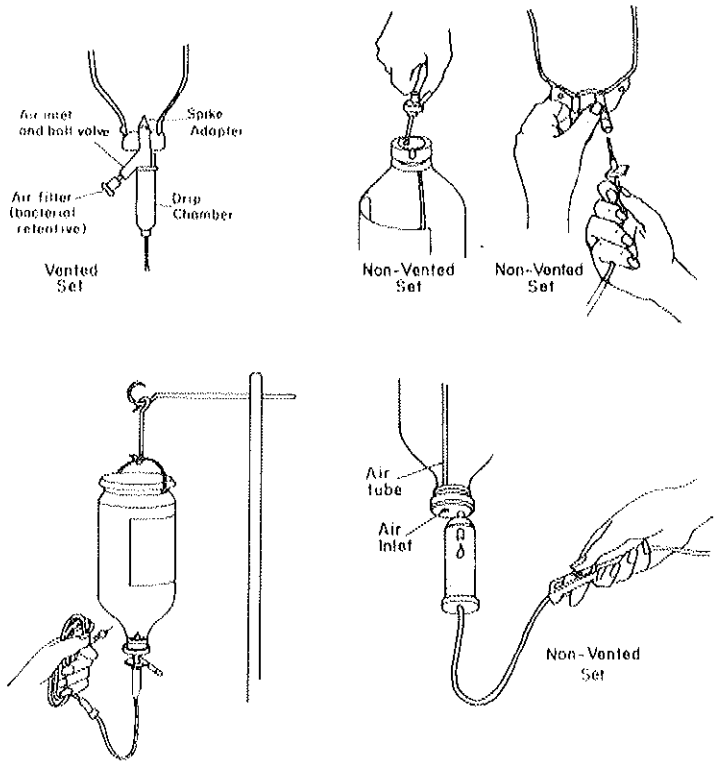


Fig 85-5. Setting up a primary IV fluid for administration.

1. The spike adapter of the administration set is inserted into the stopper or seal of the IV container. See Fig. 85-5.
2. The IV fluid is hung on a stand at bedside and air is purged from the administration set by opening the clamp until fluid comes out of needle. The tubing is then clamped off. See Fig 85-5.
3. The venipuncture is made by member of the IV team, floor nurse or physician.
4. The infusion rate is adjusted by slowly opening and closing the clamp until the desired drop rate, viewed in the drip chamber, is obtained. The usual running time is 4 to 8 hr (usually 125 mL are delivered in 1 hr). Drugs such as heparin, insulin, lidocaine or dopamine may be present in the IV drip. When potent drugs are present, the flow rates will vary depend on the clinical condition of the patient. Sets are calculated to deliver 10, 15, 20, 50 or 60 drops per mL depending on the manufacturer. See Fig 85-5.

Intermittent administration of an antibiotic and other drugs can be achieved by any of three methods: (1) direct intravenous injection (IV bolus or push), (2) addition of the drug to a predetermined volume of fluid in a volume-control device or (3) use of a second container (minibottle, minibag) with an already hanging IV fluid (piggybacking).

Direct Intravenous Injection—Small volumes (1 to 50 mL) of drugs are injected into the vein over a short period of time (1 to 5 min). The injection also can be made through a resealable gum-rubber injection site of an already hanging IV fluid. This method is suitable for a limited number of drugs but too hazardous for most drugs.

Volume-Control Method—Volume-control sets provide a means for intermittent infusion of drug solutions in precise quantities, at controlled rates of flow. These units consist of calibrated, plastic, fluid chambers placed in a direct line under an established primary IV container or more often attached to an independent fluid supply. In either case, the drug to be administered is first reconstituted if it is a sterile solid and injected into the gum-rubber injection port of the volume-control unit. It is then further diluted to 50 to 150 mL with the primary fluid or the separate fluid reservoir. Administration of the total drug-containing solution requires 30 to 60 min and produces a peak concentration in the blood followed by a valley if the dosage is discontinued. The following volume-control sets are available commercially: *Soluset*, Abbott; *Buretrol*, Baxter and *Metriset*, McGaw.

The procedure for setting up an intermittent IV infusion with a volume-control set is as follows:

Fig 85-6. Volume-control set.

1. Using aseptic technique, the spike of the volume-control set is inserted into the primary IV fluid or a separate fluid container. See Fig 85-6.
2. Air is purged from tubing of the volume-control set by opening the clamps until fluid comes through.
3. The clamp is opened above the calibrated chamber and it is filled with 25 to 50 mL fluid from the primary IV container or separate fluid container.
4. The clamp is closed above the chamber.
5. The medication is injected through the gum-rubber port of the volume-control unit.
6. The clamp above the chamber is opened to complete the dilution to the desired volume (50 to 150 mL), then closed.
7. Flow commences when the clamp below the volume-control unit is opened.

Piggyback Method—The piggyback method (Figs 85-7 and 85-8) refers to the intermittent IV drip of a second solution, the reconstituted drug, through the venipuncture site of an established primary IV system. With this setup

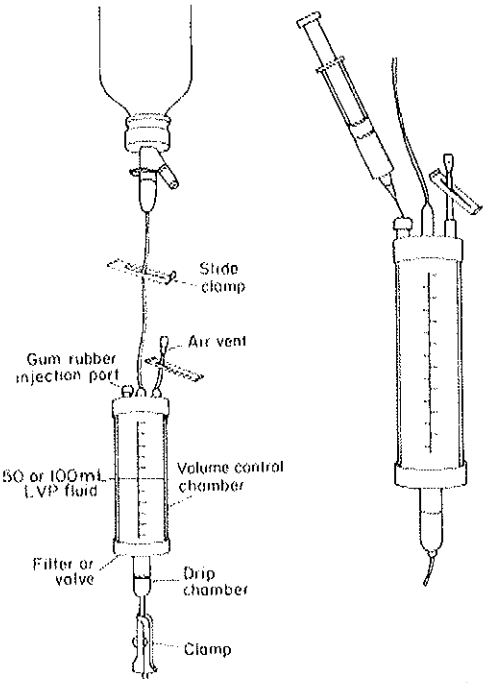


Fig 85-7. Piggyback method: the intermittent administration of a second solution through the venipuncture site of an established primary IV system.

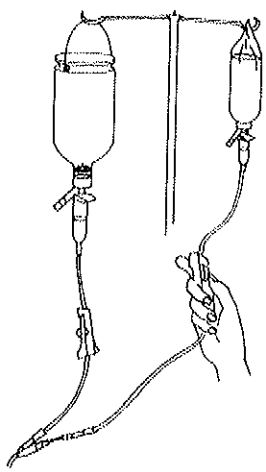


Fig 85-8. Piggyback administration setup.

the drug can be thought of as entering the vein on "top" of the primary IV fluid, hence the designation "piggyback." The piggyback technique not only eliminates the need for another venipuncture, but also achieves drug dilution and peak blood levels within a relatively short time span, usually 30 to 60 min. Drug dilution helps to reduce irritation, and early high serum levels are an important consideration in serious infection requiring aggressive drug therapy. These advantages have popularized the piggyback method of IV therapy, especially for the intermittent administration of antibiotics. In using the piggyback technique, the secondary unit is purged of air and its needle inserted into a Y-injection site of the primary set or into the injection site at the end of the primary set. The piggyback infusion is then started. Once it is completed, the primary fluid infusion will be restarted. See Fig 85-8.

Primary IV administration sets are available that have a built-in check valve for use in piggyback administration. When the piggyback is connected to one of these sets and started, the check valve automatically closes off the primary infusion. When the piggyback runs out, the check valve automatically opens, thereby restarting the primary infusion. The check valve works because of pressure differences. To achieve this difference, the primary container is hung lower than the secondary bottle by means of an extension hanger. See Fig 85-9.

Manufacturers have introduced minibottles prefilled with various antibiotic products; each container is provided with a plastic hanger for direct suspension from an IV pole as the piggyback solution is administered through the resealable gum-rubber injection site or Y-type facility of an existing IV system. Reconstitution of piggyback units requires only the addition of a small volume of compatible diluent. Since reconstitution and administration proceed from the same bottle, no drug transfer is involved, so transfer syringes and additional IV containers are not necessary. Prefilled drug containers offer significant advantages to hospitals. Time-saving, less potential for error and contamination and convenience are outstanding qualities of this type of packaging. The need exists in hospitals for these types of innovative packaging to help alleviate the critical nursing shortage and reduce the error potential. It is a significant event that drug manufacturers and intravenous fluid manufacturers have combined efforts to achieve optimal packaging for hospital use.

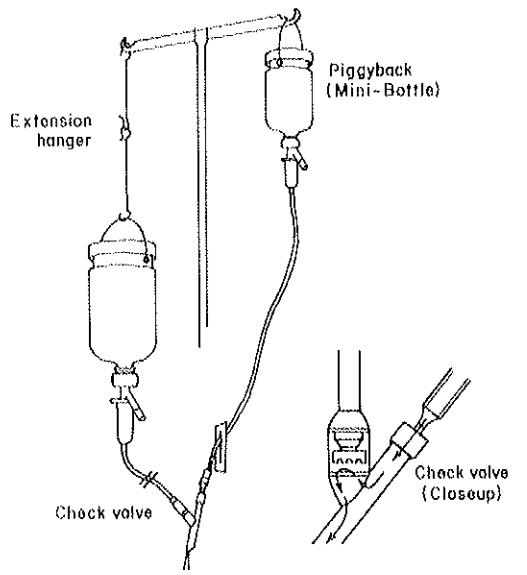


Fig 85-9. Piggyback administration setup with check valve in primary set.

Partial-fill containers available for piggybacking are 250-mL capacity infusion bottles or bags underfilled with 50 or 100 mL D5/W or normal saline. The drug to be administered first is reconstituted in its original parenteral vial and then added by needle and syringe to the partial-fill container. The needle of the piggyback delivery system is inserted into the Y-site or gum-rubber injection port of a hanging primary infusion set. Flow of the primary intravenous fluid is stopped while the drug solution in the partial-fill container is administered (30 to 60 minutes). After the drug solution has been infused totally, the primary fluid flow is reestablished. When the next dose of drug is required, the piggyback procedure is repeated, replacing the prefilled partial-fill container.

Mechanical-Electronic Infusion Devices—Gravity IV administration systems are affected by many variables which tend to alter the accuracy of the system. These include variations in the size of the drip-chamber orifice, the viscosity of the solution being administered, plastic cold flow, clamp slippage, final filters, variations in the patient's blood pressure and body movements, clot formation, pressure changes in IV containers rate of flow, temperature of the IV fluid, changes in the needle, and other factors such as kinked tubing, extravasation and changes in the height of the IV container. Flow in traditional gravity IV systems is controlled by manual clamps (either screw or roller clamps) which can provide considerable discrepancies in volume delivery. These factors have promoted the development and use of mechanical-electronic infusion devices to control more accurately the administration of IV fluids. This group of devices includes infusion controllers and infusion pumps.

Infusion controllers count drops electronically or extrude volumes of fluid mechanically and electronically. Having no moving components, controllers are less complex than pumps, being usually less expensive and having fewer maintenance problems. Infusion controllers are gravity-type systems, but the control is regulated automatically rather than manually. In addition to increasing the accuracy of delivery, electronic equipment may be able to detect infiltration of air, empty containers and excess or deficient flow.

Infusion pumps do not depend on gravity to provide the pressure required to infuse the drug. Pressure is provided by an electric pump that propels a syringe, a peristaltic or roller device or a cassette. Most pumps are volumetric in that the delivery is measured in milliliters rather than drops.

The quality of patient care has improved with the use of infusion devices. Flow rates can be maintained, therefore parenteral and enteral nutrition can be conducted safely. In addition, accurate drug therapy can be accomplished with adults and children and "runaways" of IV fluid administration can be eliminated.

Final-Filter Devices—Particulate matter in IV fluids and IV admixtures can originate from many sources. It can result from the packaging components of the IV fluid, from admixture incompatibilities, from manipulation in preparing the admixture and even from the administration set itself. Concern for particulate matter led to the design of final-filter devices for attaching to the end of the tubing of the administration set. They afford a final filtration of the IV fluid before it passes through the needle into the vein. The device consists of a plastic chamber containing a membrane or stainless-steel filter having porosities varying from 5 to 0.22 μm. Air lock can be a problem with membrane filters. When wet, membranes with a porosity of 0.22 μm and 0.45 μm are impervious to air at normal pressures and air in the system causes blockage. In order to prevent this, the filter housing must be purged completely of air prior to use. Newer designs have air eliminators. Using final-filter de-

vices increases medication cost but reduces the biological hazards associated with particulate matter.

Although considerable information is available concerning the clinical use of membrane filters in entrapping particulate matter and microorganisms, little information exists describing drug absorption by the filter. Literature on a limited number of drugs and filter materials indicates that drugs administered in low doses might present a problem with drug bonding to the filter.[1] Solutions containing minute dosages of drugs, 5 mg or less, should not be filtered until sufficient data are available to confirm insignificant absorption. Drugs not recommended to be filtered include all parenteral suspensions, blood and blood products, amphotericin B, digitoxin, insulin, intravenous fat emulsions, mithramycin, nitroglycerin and vincristine.

New IV Delivery Systems—

Frozen Premixes—Underway by Baxter is the delivery to hospitals of frozen drug products packaged in polyvinyl chloride containers. These are stored in a freezer in the hospital's pharmacy, thawed and used when needed.

Faspak/ADS-100 System—Eli Lilly supplies a non-PVC plastic piggyback container, named Faspak, which contains the dry, powdered form of certain drugs (Keflin, Kefzol, Mandol and ampicillin) which, upon reconstitution with the appropriate diluent, allows direct administration of the diluted drug. This avoids a transferring step that normally takes place when reconstituting a powdered drug. To help in the reconstitution step, a specialized dilution pump named the ADS-100 system is supplied. The package design eliminates the need for transferring between containers after reconstitution, and the Faspak acts as a final delivery container.

Abbott/ADD-Vantage System—Introduced in 1985, the Abbott ADD-Vantage system has two parts: a plastic IV bag sold by Abbott that is filled with solution and a separate glass vial of powder or liquid drug sold by a pharmaceutical manufacturer. The vial is encased by a plastic cover that is removed prior to use. The user locks the vial holding the drug into a chamber at the top of the plastic bag and mixes the drug and solution by externally removing the stopper on the vial.

Nutrimix—A Dual-Compartment container is available from Abbott. This container allows for long-term packaging of amino acids and dextrose mixtures.

IVAC-CRIS—The IVAC-Cris (Controlled-Release Infusion System) is a disposable adapter designed to infuse reconstituted injectable drugs directly from the manufacturer's single-dose vial. The CRIS adapter avoids the need to transfer drug doses to piggyback secondary containers and also eliminates the need for a secondary IV set. The adapter has a primary spike that is inserted into the IV fluid container and a secondary spike that receives the drug vial. The vial spike has two fluid paths: one admits IV fluid from the primary container into the vial; the other drains drug solution into the drip chamber of the IV set. A two-position valve allows IV fluid to flow directly from the primary container to the patient or pass through the vial to deliver the drug. A 5 μm in-line filter eliminates particulates.

To operate the CRIS adapter, the drug vial first is reconstituted with an appropriate diluent. With the valve dial in the vertical (primary) position, the spike shield is removed and the vial is attached immediately to the CRIS spike. The valve dial is then turned toward the vial, directing the flow of primary fluid into the vial of drug solution. The incoming fluid dilutes and displaces the drug solution into the drip chamber, through the primary set and into the patient. After the dose has been delivered, the vial remains on the spike until the next dose is required. Flow rate can be adjusted using a roller clamp, electronic pump or controller.

Mini-Infuser Pumps for Intermittent IV Drug Delivery—A novel concept in intermittent drug delivery, introduced several years ago, was the Bard-Harvard Mini-Infuser System. This instrument was designed for the administration of antibiotics and other medications delivered intermittently in 40 min or less. This battery-generated, lightweight instrument uses standard disposable syringes and microbore disposable extension sets. Different models are available, depending on volume-to-be-delivered selection. This instrument provides accuracy, constant flow, convenience and safety for intermittent drug delivery.

Introduced and designed for intermittent IV drug delivery, Becton Dickinson's 360 Infusor allows drug delivery intermittently over 60 min or less in a volume dilution of up to 60 mL.

Implantable Devices—The Infuse-A-Port (*Pharmacia Deltec*) was developed to satisfy the need for repeated access to the peripheral or central venous system or direct placement into an artery for regional therapy. This device may be used to withdraw blood, in addition to its use for bolus injections and short-term infusions. The Infuse-A-Port requires a special needle to allow maximum life of the self-sealing injection port.

The Infusaid Model 400 implantable drug delivery system is designed for long-term therapy in the ambulatory patient. With a 47-mL usable drug volume, it delivers a precise, continuous flow to a selected organ or site via a soft, nontraumatic, nonthrombogenic silicone rubber catheter.

Intravenous Admixtures

When one or more sterile products are added to an IV fluid for administration, the resulting combination is known as an IV admixture. To maintain the characteristics of sterile products, namely sterility, freedom from particulate matter and pyrogens, it is imperative that they be manipulated in a suitable environment using aseptic techniques.

Environment—Proper conditions for aseptic handling can be provided by laminar-flow hoods (see Chapters 78, 84). Within a laminar-flow hood, air filtered through a HEPA (high efficiency particulate air) filter moves in a parallel flow configuration at a velocity of 90 fpm. HEPA filters remove 99.97% of all particles larger than 0.3 μm. Since microbial contaminants present in air usually are found on other particulates, removal of the latter results in a flow of air free of both microbial contaminants and particulate matter. The movement of the filtered air in a laminar-flow configuration at a velocity of 90 fpm can maintain the area free of contamination. The flow of air may be in either a horizontal or vertical pattern. In the former case the HEPA filter is located at the back of the hood and the air flows to the front. In vertical flow the air passes through the HEPA filter located in the top of the cabinet and is exhausted through a grated area around the working surface of the hood. Regardless of the type of laminar air flow, the hood must be operated and maintained properly in order to achieve a satisfactory environment for the preparation of parenteral admixtures.

The hood is situated best in a clean area in which there is little traffic flow past the front of the hood. The inside of the hood is wiped down thoroughly with a suitable disinfectant and allowed to run for at least 30 min before starting manipulations. It is important to remember that the laminar-flow hood is not a means of sterilization. It only maintains an area free of microbial contaminants and particulate matter when it has been prepared, maintained and utilized properly by operators having proper aseptic techniques.

Before working in a laminar-flow hood the operator washes his hands thoroughly and scrubs them with a suitable disinfectant. Some laboratories may require gowning and

using sterile gloves. Sterile gloves can be an asset but there is always the problem that they can give the operator a false sense of security. Gloved hands can become contaminated as easily as ungloved hands. Additives and IV fluids to be used in the preparation of the admixture, along with suitable syringes, are lined up in the hood in the order they are to be used. The containers must be clean and dust-free. They are inspected for clarity and freedom from cracks. Operators are encouraged to use a lighting device for inspecting IV fluids for particulate matter and cracks. The lighting device should permit the container to be viewed against both a light and a dark background during inspection. If the IV fluid is packaged in plastic containers, pressure is applied to assure that they are sealed properly and do not leak. Some laboratories disinfect the containers prior to placing them in the hood.

In working within the hood the operator works in the center of the hood, with the space between the point of operation and the filter unobstructed. If the flow of air is blocked, the validity of the laminar flow is destroyed. Articles are arranged within the hood in a manner to prevent clean air from washing over dirty objects and contaminating other objects that must remain sterile. The working area must be at least 6 inches from the front edge of the hood. As the operator stands in front of the hood, his body acts as a barrier to the laminar air flow causing it to pass around him and create backflow patterns which can carry room air into the front of the hood.

Laminar-flow hoods must be maintained and evaluated periodically to insure that they are functioning properly. The velocity of air flow can be determined routinely using a velometer. A decrease in the air flow usually indicates a clogged HEPA filter. Some laminar-flow hoods are equipped with pressure gauges indicating pressure in the plenum behind the filter; in these hoods pressure increase also can indicate a clogged filter. Settling plates can be exposed within the hood for given periods of time to determine the presence of microbial contaminants.

The best way to determine the proper functioning of a HEPA filter is to use the dioctyl phthalate (DOP) test using the vapor at room temperature. DOP vapor (particles of ~0.3 μm) is allowed to be taken up by the hood through its intake filter. If the HEPA filter is intact and properly installed, no DOP can be detected in the filtered air stream using a smoke photometer. Certification services are available through commercial laboratories; the HEPA filters within laminar-flow hoods should be evaluated every 6 months.

Additives—The additives are injections packaged in ampuls or vials, or sterile solids; the latter are reconstituted with a suitable diluent before addition to the IV fluid. A fresh, sterile, disposable syringe is used for each additive. Before removing a measured volume from an ampul, the container is wiped with a disinfectant solution. If the ampul is scored, the top can be snapped off; if not scored, an ampul file must be used. A sterile syringe is removed from its protective wrapping. The syringe needle with its cover is separated from the syringe aseptically and may be replaced with a sterile aspirating needle. Aspirating needles usually are made from clear plastic and contain a stainless-steel or nylon filter having a porosity of 5 μm. The filter will remove glass particles and other particulates from the injection as it is drawn up from the ampul into the syringe. The aspirating needle is replaced with the regular needle. The exact volume is calibrated and the injection is ready to be added to the IV fluid (see Fig 85-10). In the case of additives packaged in multiple-dose vials, the protective cover is removed and the exposed target area of the rubber closure disinfected. A volume of air, equal to the volume of solution to be removed, is drawn up into the syringe and injected into the air space above the injection within the vial. This facilitates withdrawal of the injection. The solution is drawn into the syringe, the exact dose is measured and the injection is ready to be added to the IV fluid.

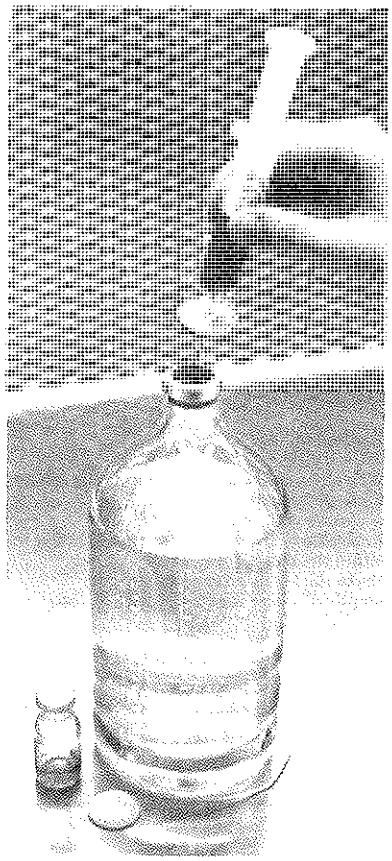


Fig 85-10. Placing an additive into an IV fluid with filtration through a membrane filter (courtesy, Millipore).

Certain injections are light-sensitive and protected against photolysis by the container packaging. The manufacturer may use amber glass, individual container wrapping or an amber plastic cover. Many hospital pharmacists use aluminum foil as a protective wrap for light-sensitive drugs during their administration.

In the case of drug substances having poor stability in aqueous solution, the drug is packaged as a sterile solid, either dry-filled or lyophilized. The diluent recommended on the labeling is used to reconstitute the powder; the proper quantity of solution then is removed for addition to the IV fluid. When large volumes of diluent are required for reconstitution, as for Keflin 4 g, a sterile needle is placed through the closure to vent the container and facilitate addition of the diluent. In order to increase the efficiency of IV admixture programs, a limited number of hospital pharmacists have found it convenient to freeze reconstituted drugs, particularly antibiotics. The stability of reconstituted drugs is somewhat limited. In some cases stability is limited to only a few hours; in many cases, however, reconstituted solutions can be frozen and thawed at the time of use. In the frozen form the stability of the antibiotic solution can be increased. In a number of instances the stability in the frozen form is known and supplied by the manufacturer. Reports have been published on the frozen stability of certain drugs. However, it is unwise to freeze drug solutions without adequate stability studies for guidance. In those cases where published information is available, close adherence must be

observed as to freezing temperature, storage conditions and packaging.

There is an increasing awareness of the potential hazard to pharmacists handling antineoplastic drugs.[2] Although the evidence is not conclusive, it appears that measures should be taken to minimize unnecessary exposure.[3] These precautions include the use of vertical laminar-flow hoods for the preparation and reconstitution of these agents, the wearing of gloves and masks by the personnel, special labeling of the containers to insure their proper handling and disposal and periodic blood studies of personnel involved in preparing admixtures of antineoplastic agents.

The procedure for placing an additive in an IV fluid will vary depending on the type of IV fluid packaging system being used by the hospital. The packaging systems have been described in Table II.

Abbott Glass Containers (Fig 85-2)

1. Remove the aluminum tear seal exposing the solid-rubber closure with a target circle in the center.
2. Wipe the closure with suitable disinfectant.
3. Insert the needle of the additive syringe through the target area. The vacuum within the bottle draws in the solution.
4. Gently shake the bottle after each addition.
5. When completed, cover the closure with a plastic protective cap if it is not to be used immediately.

Baxter and McGaw Rigid Glass Containers (Fig 85-3)

1. Remove the aluminum tear seal and the aluminum disc covering the latex diaphragm.
2. Upon exposing the latex diaphragm, note that the latex cover is drawn in over the openings in the rubber closure.
3. The larger of the two holes receives the administration set, the other is the air vent. The triangle-shaped indentation can serve as the site for injecting the additives as well as the opening for the administration set.
4. Wipe the diaphragm with a suitable disinfectant and pierce the latex cover to place additive into bottle. The vacuum within the bottle will draw additive from the syringe. Do not remove the diaphragm or the vacuum will dissipate. It will be removed at the time of administration prior to the insertion of the administration set.
5. Gently shake the bottle after each additive.
6. When completed, cover the bottle with a plastic additive cap if the administration set is not to be inserted immediately.

Baxter and Abbott Plastic Container (Fig 85-4)

1. Remove the additive port protective sleeve and rub the gum-rubber plug with a suitable disinfectant.
2. Additives are placed in container by piercing the gum-rubber cover over the additive port.
3. After each addition, milk the container to insure adequate mixing.
4. Containers do not contain a vacuum, but vacuum chambers are available for use in conjunction with the flexible plastic container.
5. Protective additive caps are available if the administration set is not inserted immediately.

McGaw Semirigid Plastic Container (Fig 85-4)

1. Remove the additive port protective covering and rub the gum-rubber plug with a suitable disinfectant.
2. Additives are placed in containers by piercing the gum-rubber over the additive port.
3. After each addition, shake the container gently to insure adequate mixing.
4. Containers do not contain a vacuum.

Parenteral Incompatibility—When one or more additives are combined with an IV fluid, their presence together may modify the inherent characteristics of the drug substances present, resulting in a parenteral incompatibility. Parenteral incompatibilities have been divided arbitrarily into three groups: physical, chemical and therapeutic. The latter is the most difficult to observe because the combination results in undesirable antagonistic or synergistic pharmacologic activity. For example, the report that penicillin or cortisone antagonizes the effect of heparin and produces a misleading picture of the anticoagulant effect of heparin represents a therapeutic incompatibility. Physical incompatibilities are observed most easily and can be detected by changes in the appearance of the admixture, such as a change in color, formation of a precipitate or evolution of a gas. Physical incompatibilities frequently can be predicted by knowing the chemical characteristics of the drugs involved. For example, the sodium salts of weak acids, such as phenytoin sodium or phenobarbital sodium, precipitate as free acids when added to intravenous fluids having an acidic pH. Calcium salts precipitate when added to an alkaline medium. Injections that require a special diluent for solubilization, such as diazepam, precipitate when added to aqueous solutions because of their low water solubility.

Decomposition of drug substances resulting from combination of parenteral dosage forms is called a chemical incompatibility, an arbitrary classification since physical incompatibilities also result from chemical changes. Most chemical incompatibilities result from hydrolysis, oxidation, reduction or complexation and can be detected only with a suitable analytic method.

An important factor in causing a parenteral incompatibility is a change in the acid-base environment.[4] The solubility and stability of a drug may vary as the pH of the solution changes. A change in the pH of the solution may be an indication in predicting an incompatibility, especially one involving drug stability, since this is not necessarily apparent physically. The effect of pH on stability is illustrated in the case of penicillin. The antibiotic remains active for 24 hr at pH 6.5, but at pH 3.5 it is destroyed in a short time. Potassium penicillin G contains a citrate buffer and is buffered at pH 6.0 to 6.5 when reconstituted with Sterile Water for Injection, Dextrose Injection or Sodium Chloride Injection. When this reconstituted solution is added to an intravenous fluid such as Dextrose Injection or Sodium Chloride Injection, the normal acid pH of the solution is buffered at pH 6.0 to 6.5, thus assuring the activity of the antibiotic.

While it may be impossible to predict and prevent all parenteral incompatibilities, their occurrence can be minimized. The IV admixture pharmacist should be cognizant of the increasing body of literature concerning parenteral incompatibilities. This includes compatibility guides published by large-volume parenteral manufacturers,[5-7] compatibility studies on individual parenteral products by the manufacturer and published with the product as part of the labeling, the study of the National Coordinating Committee on Large-Volume Parenterals,[8] reference books[9,10] and literature reports of studies with specific parenteral drugs.[11] The pharmacist should encourage the use of as few additives as possible in IV fluids since the number of potential problems increases as the number of additives increases. Physicians should be made aware of possible incompatibilities and the pharmacist can suggest alternate approaches to avoid the difficulties. In some instances, incompatibilities can be avoided by selecting another route of administration for one or more of the drugs involved.

Quality Control—Each hospital should have written procedures covering the handling and storage, use in preparing admixtures, labeling and transportation of IV fluids to the floors. In-use clarity and sterility tests should be devised to assure that IV admixtures retain the characteristics of sterility and freedom from particulate matter. Training and monitoring personnel involved in preparation of IV admixtures should be done on a regular basis.[12] The efforts of the hospital pharmacy should be no less than those of the industry in following Current Good Manufacturing Practice to assure the safety and efficacy of these compounded medications.

Total Parenteral Nutrition

Intravenous administration of calories, nitrogen and other nutrients in sufficient quantities to achieve tissue synthesis and anabolism is called total parenteral nutrition (TPN).[13]

Originally, the term hyperalimentation was used to describe the procedure, but it is being replaced by TPN, the latter being more descriptive for the technique.

The normal caloric requirement for an adult is approximately 2500 per day. If these were to be provided totally by D5/W, approximately 15 L would be required. Each liter contains 50 g dextrose, equivalent to 170 calories. However, it is only possible to administer 3 or 4 L per day without causing fluid overload. To reduce this fluid volume the concentration of dextrose would have to be increased. By increasing the dextrose to 25%, it is possible to administer five times the calories in one-fifth the volume. D25/W is hypertonic and cannot be administered in large amounts into a peripheral vein without sclerosing the vein.

Dudrick developed the technique for administering fluids for TPN by way of the subclavian vein into the superior vena cava where the solution is rapidly diluted by the large volume of blood available, thus minimizing the hypertonicity of the solution. For administration of the TPN fluids, a catheter is inserted and retained in place in the subclavian vein. TPN is indicated in patients who are unable to ingest food due to carcinoma or extensive burns; patients who refuse to eat, as in the case of depressed geriatrics or young patients suffering from anorexia nervosa and surgical patients who should not be fed orally.

The preferred source for calories in TPN fluids is the carbohydrate dextrose. Both fat emulsions and alcohol are caloric sources, but they are not used in TPN fluids. In IV fluid kits commercially available for the preparation of TPN solutions, D50/W is provided. On dilution with amino acid injection, the resulting dextrose concentration is approximately 25%. It is this concentration that is administered.

The source of nitrogen in TPN fluids is crystalline amino acids (*Aminosyn*, Abbott; *FreAmine III*, McGaw; *Travasol*, Travenol). The crystalline amino acid injections contain all the essential and nonessential amino acids in the L-form. For optimum utilization of amino acids and for promoting tissue regeneration, the nitrogen-to-calorie ratio should be 1:150. Calories are needed to provide energy for the metabolism of nitrogen.

Electrolyte requirements vary with the individual patient. The electrolytes present in Amino Acid Injection are given

Table III—Typical IV Orders (Parenteral Prescriptions)

Prescription	Comment
1. ℞ NSS 1000 mL 125 mL/hr	Sodium Chloride Injection (Normal Saline Solution) 1000 mL, is to be administered at a flow rate of 125 mL per hr. It will require approximately 8 hr.
2. ℞ 1000 D5 + NSS + Vits 12 hr	Dextrose Injection 5%, 1000 mL, containing 0.9% sodium chloride and container of vitamin B complex with vitamin C is to be administered over a 12-hr period.
3. ℞ 500 D5 + ½NSS KVO	Dextrose Injection 5%, 500 mL, containing 0.45% sodium chloride is to be administered at a flow rate to keep the vein open (KVO). The flow rate will be approximately 10 mL per 1 hr.
4. ℞ 1000 cc D5 + ½NSS Add 1 amp Vits to each + 100 mg Thiamine Each to run 6 hr	Dextrose Injection 5%, 1000 mL, containing 0.45% sodium chloride, the contents of one ampul vitamin B complex with vitamin C and sufficient volume of Thiamine Hydrochloride Injection to give 100 mg thiamine, is to be administered over a 6-hr period (approximately 170 mL per hr). Additional orders of the same can be anticipated.
5. ℞ 1000 cc D5 + ½NSS + 20 mEq KCl	Dextrose Injection 5%, 1000 mL, is to be provided containing 0.45% sodium chloride and 20 mEq potassium chloride.
6. ℞ 1000 Hyperal + 10 NaCl + 10 KCl + 5 $MgSO_4$ + 10 Insulin	One L of the hospital's basic TPN solution is to be provided with the addition of 10 mEq sodium chloride, 10 mEq potassium chloride, 5 mEq magnesium sulfate and 10 units regular zinc insulin.
7. ℞ 1000 cc Hyperal (FreAmine) + 40 mEq $NaHCO_3$ + 30 mEq KCl + Vits + 5U Reg Insulin to run 80 cc/hr	One L of the basic TPN solution, FreAmine II, is to be provided with the addition of 40 mEq $NaHCO_3$, 30 mEq potassium chloride, the contents of one container vitamin B complex with vitamin C plus 5 units of regular zinc insulin. It is to be administered at the flow rate of 80 mL per hr (approximately 12 hr).
8. ℞ 1000 Hyperal + 40 mEq NaCl + 10 KCl + 10 Insulin + 10 Cal Gluconate	One L of the hospital's basic TPN solution is to be provided with the addition of 40 mEq sodium chloride, 10 mEq potassium chloride, 10 units regular zinc insulin and 10 mL Calcium Gluconate Injection.
9. ℞ Keflin 2 g + 100 mL D_5W q 6 hr	Cephalothin, 2 g, is reconstituted with Sterile Water for Injection and added to a minibottle containing 100 mL Dextrose Injection 5%. This dose is given every 6 hr using a piggyback technique with a flow rate requiring 30 to 60 min for delivery.
10. ℞ Gentamicin 80 mg IVPB q 8 hr	Gentamicin, 80 mg, is added to a minibottle containing 100 mL Dextrose Injection 5%. This dose is given every 8 hr using the piggyback technique (IVPB) with a flow rate requiring at least 80 min (not less than 1 mg per min).

on the label and must be taken into consideration in determining the quantities to be added. Usual electrolyte concentrations are required to fall within the following ranges: sodium, 100–120 mEq; potassium, 80–120 mEq; magnesium, 8–16 mEq; calcium, 5–10 mEq; chloride, 100–120 mEq and phosphate, 40–60 mEq. It is better to keep a 1:1 ratio between sodium and chloride ions. In adding potassium, the acetate salt is preferred to the chloride. If the combination of calcium and phosphate ions exceeds 20 mEq, precipitation occurs.

In addition to the electrolytes, the daily requirement for both water-soluble and fat-soluble vitamins may be added, usually in the form of a multivitamin infusion concentrate. Iron, should be administered separately from the TPN fluids. Trace elements such as zinc, copper, manganese and iodide are a concern only in long-term cases and can be added when required.

The Parenteral Prescription

The physician writes an admixture order or parenteral prescription on a physician's order-form located on the patient's chart. A copy of the order is sent to the pharmacy for compounding. It includes the patient's name, room number, the intravenous fluid wanted, additives and their concentrations, rate of flow, starting time and length of therapy. The order is taken by the technician, nurse or pharmacist to the pharmacy. Orders may be telephoned to the pharmacy; verification with the original order is made on delivery of the admixture. IV orders usually are written for a 24-hour therapy period; the patient's chart is reviewed and new orders are written on a daily basis. The order may be for multiple containers, in which case the containers are numbered consecutively. Unlike the extemporaneously compounded prescription, additives are added without regard to final volume of IV fluid. The prescription is checked for proper dose, compatibility, drug allergies and stability. Additives usually are given an expiration period of 24 hours from the time of preparation. Drugs such as ampicillin may require shorter expiration periods.

The clerical work for the admixture is prepared. This includes typing of the label and the preparation of the profile worksheet. The profile sheet is filed so that the pharmacist will be alerted when subsequent containers are due for preparation. Charging the patient's account can be done from the profile worksheet. The label includes the patient's name, room number, bottle number, preparation date, expiration time and date, intravenous fluid and quantity, additives and quantities, total time for infusion, the milliliters per hour or drops per minute and space for the name of the nurse who hangs the container. The label will be affixed to the container upside down in order that it can be read when hung.

The admixture is prepared by the pharmacist or a supervised technician. In handling sterile products, aseptic techniques as discussed previously must be observed. When completed, a plastic additive cap is affixed before delivery to the floor. The label is applied and checked with the original order. The empty additive containers are checked to confirm the additives present. The admixture is inspected for any color change or particulate matter.

The completed admixture is delivered to the floor. If it is not to be infused immediately (within 1 hour), it is stored under refrigeration; if refrigerated, it must be used within 24 hours. The nurse checks for accuracy of patient's name, drug and concentration, IV fluid, expiration date, time started and clarity. The infusion of admixtures may run ahead or behind schedule, necessitating that the pharmacist modify the preparation of continued orders. Examples of IV orders are shown in Table III.

References

1. Turco SJ: Drug absorption to membrane filters. *Am J IV Ther Clin Nutr 9*: 6, 1982.
2. Zimmerman PF, Randolph KL, Barkley EW, Gallelli JF: Recommendations for the safe handling of injectable antineoplastic drug products. *Am J Hosp Pharm 38*: 1693, 1981.
3. Gallelli JF: Evaluating the potential hazard of handling antineoplastic drugs. *Am J Hosp Pharm 39*: 1877, 1982.
4. Newton DW: Physicochemical determinants of incompatibility and instability of injectable drug solutions and admixtures. *Am J Hosp Pharm 35*: 1213, 1978.
5. King JC: *Guide to Parenteral Admixtures*. Cutter Laboratories, Berkeley CA, 1987.
6. Shoup LK, Goodwin NH: *Implementation Guide—Centralized Admixture Program*. Travenol Laboratories, Morton Grove IL, 1977.
7. *Good IV Procedures Manual*. Abbott Laboratories, North Chicago IL, 1979.
8. Bergman HD: Incompatibilities in large-volume parenterals. *Drug Intell Clin Pharm 11*: 345, 1977.
9. Trissel LA: *Parenteral Drug Information Guide*, Am Soc Hosp Pharm, Washington DC, 1974.
10. Trissel LA: *Handbook on Injectable Drugs*, 4th Ed, Am Soc Hosp Pharm, Washington DC, 1987.
11. Kobayashi NH, King JC: Compatibility of common additives in protein hydrolysate/dextrose solutions. *Am J Hosp Pharm 34*: 589, 1977.
12. Sanders SJ, Mabadeje SA, Avis KE, Cruze CA, Martinez DR: Evaluation of compounding accuracy and aseptic techniques for intravenous admixtures. *Am J Hosp Pharm 35*: 531, 1978.
13. Dudrick SJ, Rhoads JE: Total intravenous feeding. *Sci Am 226*: 73, 1972.

Bibliography

Turco SJ, King RE: *Sterile Dosage Forms: Their Preparation and Clinical Applications*, 3rd, Lea & Febiger, Philadelphia, 1987.

Trissel LA: *Handbook on Injectable Drugs*, 4th ed, Am Soc Hosp Pharm, Washington DC, 1987.

National Coordinating Committee on Large-Volume Parenterals: Recommended methods for compounding intravenous admixtures in hospitals. *Am J Hosp Pharm 32*: 261, 1975.

Avis, KE, Akers MJ: *Sterile Preparation for the Hospital Pharmacist*, Ann Arbor Sci Publ, Ann Arbor MI, 1981.

CHAPTER 87

Medicated Applications

Lawrence H Block, PhD
Professor of Pharmaceutics
Duquesne University School of Pharmacy
Pittsburgh, PA 15282

The application of medicinal substances to the skin or various body orifices is a concept doubtless as old as humanity. The papyrus records of ancient Egypt describe a variety of such medications for external use. Galen described the use in Roman times of a forerunner to today's vanishing creams.

Medications are applied in a variety of forms reflecting the ingenuity and scientific imagination of pharmacists through the centuries. New modes of drug delivery have been developed to remedy the shortcomings of earlier vehicles or, more recently, to optimize drug delivery. Conversely, some external medications have fallen into disuse because of changes in the practice of medicine.

Medications are applied to the skin or inserted into body orifices in liquid, semisolid or solid form. Ophthalmic ointments and topical aerosol products will not be discussed in this chapter. Ophthalmic use imposes unusual particle size, viscosity and sterility specifications that require separate, detailed discussion (see Chapter 86). The complexity of pharmaceutical aerosol systems necessitates their inclusion elsewhere (see Chapter 92).

Epidermal and Transdermal Drug Delivery

The Skin

The skin often has been referred to as the largest of the body organs: an average adult's skin has a surface area of about 2 m^2. It is probably the heaviest organ of the body. Its accessibility and the opportunity it affords to maintain applied preparations intact for a prolonged time have resulted in its increasing use as a route of drug administration, whether for local, regional or systemic effects.

Anatomically, human skin may be described as a stratified organ with three distinct tissue layers: the epidermis, the dermis and the subcutaneous fat layer (Fig 87-1).

Epidermis, the outermost skin layer, comprises stratified squamous epithelial cells. Keratinized, flattened remnants of these actively dividing epidermal cells accumulate at the skin surface as a relatively thin region (about 10 μm thick) termed the stratum corneum, or horny layer. The horny layer is itself lamellar with the keratinized cells overlapping one another and compressed into about 15 layers. The region behaves as a tough but flexible coherent membrane. The stratum corneum also is markedly hygroscopic—far more so than other keratinous materials such as hair or nails. Immersed in water the isolated stratum corneum swells to about three times its original thickness, absorbing about four to five times its weight in water in the process. The stratum corneum functions as a protective physical and chemical barrier and is only slightly permeable to water. It retards water loss from underlying tissues, minimizes ultraviolet light penetration and limits the entrance of microorganisms, medications and toxic substances from without. The stratum corneum is abraded continuously. Thus, it tends to be thicker in regions more subject to abrasion or the bearing of weight. Its regeneration is provided by rapid cell division in the basal cell layer of the epidermis. Migration or displacement of dividing cells towards the skin surface is accompanied by differentiation of the epidermal cells into layers of flat, laminated plates, as noted above. An acidic film (pH ranging between 4.0 and 6.5, depending on the area tested) made up of emulsified lipids covers the surface of the stratum corneum.

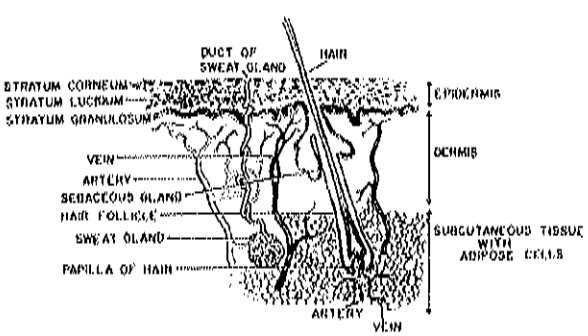


Fig 87-1. Vertical section of human skin.

The dermis apparently is a gel structure involving a fibrous protein matrix embedded in an amorphous, colloidal, ground substance. Protein, including collagen and elastin fibers, is oriented approximately parallel to the epidermis. The dermis supports and interacts with the epidermis facilitating its conformation to underlying muscles and bones. Blood vessels, lymphatics and nerves are found within the dermis, though only nerve fibers reach beyond the dermal ridges or papillae into the germinative region of the epidermis. Sweat glands and hair follicles extending from the dermis through the epidermis provide discontinuities in an otherwise uniform integument.

The subcutaneous fat layer serves as a cushion for the dermis and epidermis. Collagenous fibers from the dermis thread between the accumulations of fat cells providing a connection between the superficial skin layers and the subcutaneous layer.

Hair Follicles and Sweat Glands—Human skin is sprinkled liberally with surface openings extending well into the dermis. Hair follicles, together with the sebaceous glands that empty into the follicles, make up the pilosebaceous unit. Apocrine and eccrine sweat glands add to the total.

Pilosebaceous Unit—Human hair consists of compacted keratinized cells formed by follicles. Sebaceous glands empty into the follicle sites to form the pilosebaceous unit. The hair follicles are surrounded by sensory nerves; thus, an important function of human hair is sensory. Human hair varies enormously within the same individual, even within the same specific body area. Individual hairs can vary in microscopic appearance, diameter, cuticle appearance and even presence or absence of medulla.

Sebaceous glands are similar anatomically and functionally but vary in size and activity according to location. Popu-

Table I—Composition of Sebum

Constituents	Percent w/w
Triglycerides	57.5
Wax Esters	26.0
Squalene	12.0
Cholesterol Esters	3.0
Cholesterol	1.5

lation in the scalp, face and anogenital areas may vary from 400 to 900/cm^2. Fewer than 100/cm^2 are found in other areas. Sebaceous glands are richly supplied with blood vessels.

Sebaceous cells synthesize and accumulate lipid droplets. This accumulation results in enlarged cells which fragment to form sebum. Sebum is made up of a mixture of lipids, approximately as shown in Table I.

The sebaceous gland, containing sebum, cell debris and microorganisms such as *Propionibacterium acnes*, is connected to the pilosebaceous canal by a duct of squamous epithelium. When access to the surface is blocked and bacteria multiply, the result is the comedo of acne.

Sebum presumably functions as an emollient, although Kligman once stated it was useless. Montagna suggests that sebum functions as a pheromone to provide the human with a distinctive aroma.

Sweat Glands—Sweat glands are classified as apocrine and eccrine. Apocrine glands are secretory but are not necessarily responsive to thermal stimulation. Such glands do not produce sweat in the normal sense of the word. Apocrine glands, however, often are associated with eccrine sweat glands particularly in the axilla.

Eccrine sweat glands are coiled secretory glands, equipped with a blood supply, extending from the dermis to the epidermal surface. Eccrine sweat glands function to regulate heat exchange in man. As such, they are indispensable to survival.

About 3 million eccrine glands are thought to be distributed over the human body. Distribution varies from less than 100 to more than 300/cm^2. Gland counts after thermal stimulation do not always agree with anatomical counts.

Drug Effects and the Extent of Percutaneous Drug Delivery

Drugs are applied to the skin to elicit one or more of four general effects: an effect on the skin surface, an effect within the stratum corneum, a more deep-seated effect requiring penetration into the epidermis and dermis or a systemic effect resulting from delivery of sufficient drug through the epidermis and the dermis to the vasculature to produce therapeutic systemic concentrations.

Surface Effects—An activity on the skin surface may be in the form of a film, an action against surface microorganisms or a cleansing effect. Film formation on the skin surface may be protective, *eg*, a zinc oxide cream or a sunscreen. Films may be somewhat occlusive and provide a moisturizing effect by diminishing loss of moisture from the skin surface. In such instances the film or film formation *per se* fulfils the objective of product design. The action of antimicrobials against surface flora requires more than simple delivery to the site. The vehicle must facilitate contact between the surface organisms and the active ingredient. Skin cleansers employ soaps or surfactants to facilitate the removal of superficial soil.

Stratum Corneum Effects—Drug effects within the stratum corneum are seen with certain sunscreens; *p*-aminobenzoic acid is an example of a sunscreening agent which both penetrates and is substantive to stratum corneum cells. Skin moisturization takes place within the stratum corneum. The dry outer cells are hydrated by surface films. The increased moisture results in an apparent softening of the skin. Keratolytic agents, such as salicylic acid, act within the stratum corneum to cause a breakup or sloughing of stratum corneum cell aggregates. This is particularly important in conditions of abnormal stratum corneum such as psoriasis, a disease characterized by thickened scaly plaques.

The stratum corneum also may serve as a *reservoir phase* or depot wherein topically applied drug accumulates due to partitioning into or binding with skin components. This interaction can limit the subsequent migration of the penetrant unless the interaction capacity of the stratum corneum is surpassed by providing excess drug. Examples of drugs which exhibit significant skin interaction include benzocaine, scopolamine and corticosteroids.

Epidermal, Dermal, Local and Systemic Effects—The penetration of a drug into the viable epidermis and dermis may be difficult to achieve, as noted above. But, once transepidermal permeation has occurred, the continued diffusion of drug into the dermis is likely to result in drug transfer into the microcirculation of the dermis and then into general circulation. Nonetheless, it is possible to formulate drug delivery systems which provide substantial localized delivery without achieving correspondingly high systemic concentrations. Limited studies in man of topical triethanolamine salicylate, minoxidil and retinoids demonstrate the potential of this approach.

Unwanted systemic effects stemming from the inadvertent transdermal penetration of drugs have been reported for a wide variety of compounds (eg, hexachlorophene, lindane, corticosteroids) over the years. With the commercial introduction of transdermal drug delivery systems for scopolamine, nitroglycerin, clonidine and 17β-estradiol, transdermal penetration is being regarded increasingly as an opportunity rather than a nuisance.

Percutaneous Absorption

Percutaneous absorption involves the transfer of drug from the skin surface into the stratum corneum, under the aegis of a concentration gradient, and its subsequent diffusion through the stratum corneum and underlying epidermis, through the dermis and into the microcirculation. The skin behaves as a passive barrier to diffusing molecules. Evidence for this includes the fact that the impermeability of the skin persists long after the skin has been excised. Furthermore, Fick's Law is obeyed in the vast majority of instances.

Molecular penetration through the various regions of the skin is limited by the diffusional resistances encountered. The total diffusional resistance (R_{skin}) to permeation through the skin has been described by Chien as

$$R_{skin} = R_{sc} + R_e + R_{pd}$$

where R is the diffusional resistance and the subscripts *sc*, *e* and *pd* refer to the stratum corneum, epidermis and papillary layer of the dermis, respectively. In addition, resistance to transfer into the microvasculature limits the systemic delivery of drug.

By and large, the greatest resistance to penetration is met in the stratum corneum, ie, diffusion through the stratum corneum tends to be the rate-limiting step in percutaneous absorption.

The role of hair follicles and sweat glands must be considered; however, as a general rule their effect is minimized by the relatively small fractional areas occupied by these appendages. In the very early stages of absorption, transit through the appendages may be comparatively large, partic-

ularly for lipid-soluble molecules and those whose permeation through the stratum corneum is relatively low.

The stratum corneum can be regarded as a passive diffusion membrane but not an inert system; it often has an affinity for the applied substance. The adsorption isotherm is frequently linear in dilute concentration ranges. The correlation between external and surface concentrations is given in terms of the solvent membrane distribution coefficient K_m. The integrated form of Fick's Law is given as

$$J_s = \frac{K_m D C_s}{\delta}$$

and

$$K_p = \frac{K_m D}{\delta}$$

where K_p is the permeability coefficient, J_s is the steady state flux of solute, C_s is the concentration difference of solute across membrane, δ is the membrane thickness,

$$K_m \text{ is the } \frac{\text{solute sorbed per cc of tissue}}{\text{solute in solution per cc solvent}} = \frac{C_m}{C_s}, \text{ and}$$

D is the average membrane diffusion coefficient for solute.

Permeability experiments have shown that the hydrated stratum corneum has an affinity for both lipophilic and hydrophilic compounds. The bifunctional solubility arises from the filament-matrix ultrastructure of the keratin, which allows aqueous and lipid regions to coexist. Thus, attempts to predict permeability constants from oil:water or solvent:water partition coefficients have had limited success.

The effect of regional variation on skin permeability can be marked. Kligman suggests that two species of horny layer be recognized: the palms and soles, adapted for weight-bearing and friction; and the body horny layer, adapted for flexibility, impermeability and sensory discrimination.

Overall, data suggest the following order for diffusion of simple molecules through the skin: plantar > palmar > dorsum of hand > scrotal and postauricular > axillary > scalp > arms, legs, trunk. Electrolytes in solution penetrate the skin poorly. Ionization of a weak electrolyte substantially reduces its permeability, eg, sodium salicylate permeates poorly compared with salicylic acid. Nonetheless, the development of iontophoretic devices in recent years may minimize this problem with ionic penetrants.

In Vitro and In Vivo Studies

Classically, percutaneous absorption has been studied *in vivo* using radioactively labeled compounds or by *in vitro* techniques using excised human skin. A diffusion cell frequently used for *in vitro* experiments is shown in Fig 87-2.[1] In this system the intact skin or the epidermis is treated as a semipermeable membrane separating two fluid media. The transport rate of a particular drug is evaluated by introducing the drug in solution on the stratum corneum side of the "membrane," then measuring penetration by periodic sampling and analysis of the fluid across the skin membrane.

More recently investigators have recognized that transport across an immersed, fully hydrated stratum corneum may not represent the absorption system or rate observed in *in vivo* studies. Percutaneous absorption across a fully hydrated stratum corneum may be an exaggeration. It may be more representative of enhanced absorption that is seen after *in vivo* skin is hydrated by occlusive wrapping.

Using separated epidermal skin mounted in diffusion cells, Scheuplein and Ross[2] varied the atmosphere above the skin strip by use of Drierite to simulate dry conditions and wetted paper strips to simulate the effect of occlusion and observed marked reduction in penetration of cortisone under dry conditions but greatly enhanced penetration on humidifying the stratum corneum (see Fig 87-3).[2]

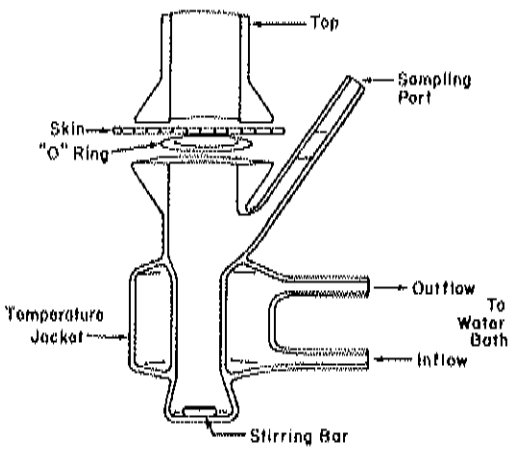


Fig 87-2. Schematic representation of diffusion cell. Top is open to ambient laboratory environment.[1]

The studies of Scheuplein and Ross,[2] and of Franz,[1] demonstrate that *in vitro* studies of percutaneous absorption under controlled conditions are relevant to *in vivo* drug penetration. As stated by Franz, "whenever a question is asked requiring only a qualitative or directional answer, the *in vitro* technique appears perfectly adequate."

Relevance of Animal Studies

Any evaluation of a study of percutaneous absorption in animals must take cognizance of species variation. Just as percutaneous absorption in man will vary considerably with skin site, so will absorption in various animal species. Bartek *et al*[3] investigated percutaneous absorption and found a decreasing order of permeability, thus, rabbit > rat > swine > man. They studied the *in vivo* absorption of radioactively labeled haloprogin, *N*-acetylcysteine, testosterone, caffeine and butter yellow; their results with testosterone, shown in Fig 87-4, illustrate the penetration differences observed with different animal skins.

Subsequently, using a similar *in vivo* technique, Wester

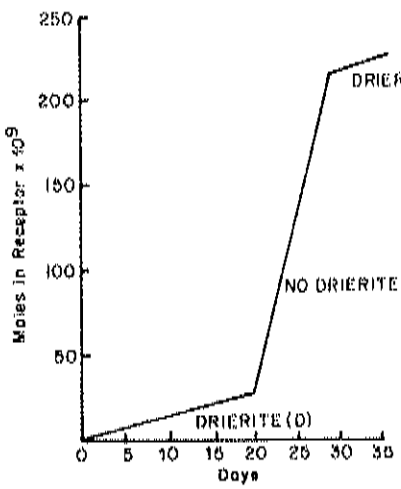


Fig 87-3. Change in cortisone penetration by alternately drying (D) and humidifying (W) the stratum corneum.[2]

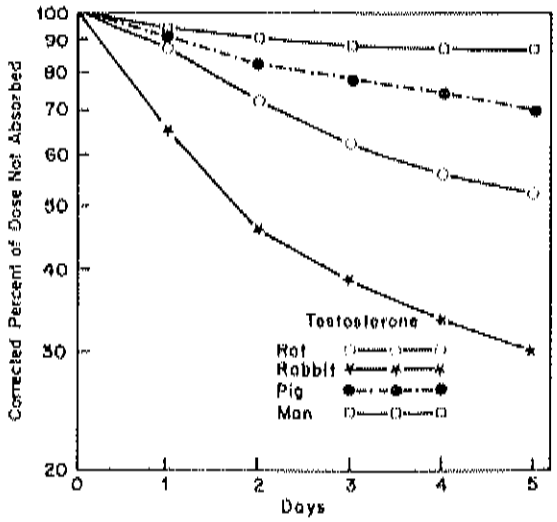


Fig 07-4. Percutaneous absorption of testosterone in rats, rabbits, swine and man for 5 days after application.[4]

and Maibach[5] investigated the percutaneous absorption of benzoic acid, hydrocortisone and testosterone in the rhesus monkey. Radioactively tagged compounds were applied to the ventral surface of the forearm, and absorption was quantified on the basis of radioactivity excreted in the urine for five days following application. The investigators concluded that the percutaneous penetration of these compounds in the rhesus monkey is similar to that in man, and regarded the data as encouraging because of the similarity.

Stoughton[6] performed *in vitro* studies using animal skins. Using a variety of compounds and a diffusion-cell apparatus he concluded that the skin of the hairless mouse or the baby rat is useful for screening for absorption or epidermal response.

It should be stressed again that percutaneous absorption studies in animals, either *in vivo* or *in vitro*, only can be useful approximations of activity in man. The effect of species variation, site variability (about which little is known in animals), skin condition, experimental variables and, of major importance, the vehicle, must be kept in mind.

Drug Testing in Animals

Drug testing in animals is a characteristic of new-drug development, and the testing of dermatological products or drugs in animals is no exception. Such testing typically may take three forms: Animals may be used to estimate the safety of a drug product or substance; animal skin may be substituted for human skin for a specific measurement, eg, percutaneous absorption; animal skin may be used as a disease model to simulate an equivalent human condition.

Animals have been used to detect contact sensitization, measure antimitotic drug activity, measure phototoxicity and evaluate the comedogenic and comedolytic potential of substances. In each of these test procedures, be it a safety test or assay model, the animal is considered a substitute for man. It is, therefore, important to realize that the animal is not man, even though man is the ultimate test animal. Animal-testing presents the investigator with unique advantages; lack of appreciation of the variables involved can destroy these advantages.

Mershon and Callahan[7] recorded and illustrated the considerations involved in selecting an animal test model. They interpreted the rabbit irritancy data of several investigators, and impressively visualized different possible interpretations of the differing response between rabbit and man.

Table II—Relative Potency of Anti-Inflammatory Agents[9]

Compound	Topical anti-inflammatory potency: Rat-ear edema assay	Topical anti-inflammatory potency: Human vasoconstrictor assay
Dexamethasone	73.2 (49.4–110)	10–20
Dexamethasone 21-acetate	117.3 (85.9–106)	10–20
Prednisolone	2.44 (1.54–7.76)	1–2
Prednisolone 21-acetate	5.43 (4.05–7.70)	3
Betamethasone	97.3 (16.7–141)	3–5
Betamethasone 21-acetate	1072.0 (876–1179)	18–33
Fluorometholone	138.3 (57.9–333)	30–40
Fluorometholone acetate	219.5 (9.15–536)	
Fluprednisolone	31.8 (13.3–76.1)	4–6
Fluprednisolone acetate	61.3 (25.6–147)	
Hydrocortisone	1	1

( ) = 95% confidence limits

While the ultimate system for establishing therapeutic efficacy is man, there are specific animal test models that are recognized to be valuable as prehuman-use screens predictive of drug activity in humans. For example, the rat-ear assay and the granuloma-pouch procedure in rats are recognized procedures for the estimation of steroid anti-inflammatory activity.

Lorenzetti[8] tabulated the potency of various topical steroids, comparing the rat-ear-edema assay with potency measured in humans using the vasoconstrictor procedure of Stoughton and McKenzie; the results are given in Table II.[9] Animal assay models of this kind, particularly the steroid anti-inflammatory assays, are most useful as preliminary activity screens. The simplicity, safety and reproducibility of the vasoconstrictor assay in humans recommend it over any corresponding animal procedure.

In Numero Models

In recent years, *in numero* modeling or computer simulation of percutaneous absorption has been advocated as a link between *in vitro* and *in vivo* studies. A number of relatively simplistic dermatopharmacokinetic models have been developed that do provide the formulator with some insight into transdermal drug delivery, in spite of the biological and physicochemical complexity of drug transport into and through the skin. By and large, these models are analogous to the classical pharmacokinetic models which have been employed to assess *in vivo* drug uptake and disposition. Some of the dermatopharmacokinetic models proposed differ from more classically oriented models in that drug transport in the vehicle and in the epidermis, particularly the stratum corneum, is modeled in accordance with Fickian diffusion. Thus, the formulator can anticipate the effect of variables such as the thickness of the applied (vehicle) phase, alterations in drug partitioning between the vehicle and the stratum corneum and the frequency of reapplication on the overall appearance of drug systemically as a function of time following topical application.

Dosage-Form Design

More than 35 years ago Lane and Blank pointed out that sufficient thought rarely is given to the function which the vehicle performs and to the physicochemical characteristics of the base. These investigators were not discussing optimization of drug activity in today's meaning of the term. They emphasized that the type of skin, application site, lesion type and physicochemical action of the base are important considerations.

In many (if not most) clinical situations the rate-limiting step is penetration of the drug across the skin barrier, ie,

percutaneous penetration through the skin alone. Diffusion of the drug from its vehicle, although dependent on the same diffusion parameters, should not be unknowingly the rate-limiting step in percutaneous absorption. Such a rate limitation or control may, of course, be an objective and the end point of specific drug optimization, but inappropriate formulation can reduce substantially the effectiveness of a topical drug substance.

In the formulation of a vehicle for topical drug application many factors must be considered. Drug stability, specific product use, site of application and product type must be combined in a dosage form which will readily release the drug when placed in contact with the skin. Further, the release characteristics of the vehicle are dependent on the physical-chemical properties of the specific drug substance to be delivered to the skin. A vehicle optimized for delivery of hydrocortisone may be quite inappropriate for delivery of a different steroid.

T Higuchi discussed (1960 to 1961) equations describing the rate of release of solid drugs suspended in ointment bases. Ostrenga *et al*, in a series of publications, discussed the significance of vehicle composition on the percutaneous absorption of fluocinolone acetonide and fluocinolone acetonide 21-acetate (fluocinonide) (see Fig 87-5).[10] These investigators used propylene glycol/isopropyl myristate partition coefficients, *in vitro* (human) skin penetration and finally *in vivo* vasoconstrictor studies to evaluate formulation variables. They concluded that

"In general, an efficacious topical gel preparation is one in which (a) the concentration of diffusible drug in the vehicle for a given labeled strength is optimized by ensuring that all of the drug is in solution, (b) the minimum amount of solvent is used to dissolve the drug completely and yet maintain a favorable partition coefficient and (c) the vehicle components affect the permeability of the stratum corneum in a favorable manner."

The effect of propylene glycol concentration on *in vivo* vasoconstrictor activity is illustrated strikingly in Fig 87-5, taken from Ostrenga *et al*.[10]

Experimental work of the kind described by Ostrenga *et al*[10] provides a means of optimizing drug release from a vehicle and penetration of the drug into the skin. This is a beginning. The formulator must proceed to develop a total composition in which the drug is stable and causes no irritation to sensitive skin areas. Safety, stability and effective preservative efficacy must be combined with optimum drug delivery in the total formulation.

Optimization of drugs other than steroids may be approached by direct *in vivo* assays. Layers of the stratum corneum can be removed or stripped successively away by the repeated application and removal of cellulose adhesive tape strips. The penetration into the skin, as well as the effect of additives on *p*-aminobenzoic acid, were studied by Lorenzetti through analysis of individual skin strips. The results provided a profile of skin penetration and visualized the effect of additives. Similar experiments have been carried out using benzoyl peroxide. Penetration *per se*, as well as the effect of additives, can be measured by chemical analysis of individual tape strips following application of a specific quantity of drug or drug product.

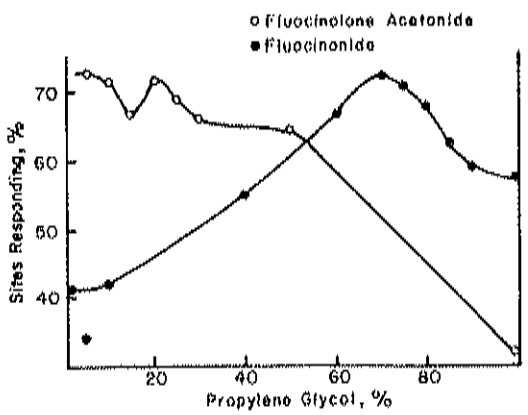


Fig 87-5. *In vivo* response as a function of vehicle composition (24-hour vasoconstriction).[10]

Factors Affecting Drug Absorption

In the foregoing it has been seen that drug-release from its vehicle is a function of concentration, solubility in the vehicle and partition coefficient between the vehicle and the receptor site. Percutaneous absorption of a drug also can be enhanced by the use of occlusive techniques or by the use of so-called penetration enhancers.

Skin Hydration and Temperature—Occluding the skin with wraps or impermeable plastic film such as Saran Wrap prevents the loss of surface water from the skin. Since water is absorbed readily by the protein components of the skin the occlusive wrap causes greatly increased levels of hydration in the stratum corneum. The concomitant swelling of the horny layer ostensibly decreases protein network density and the diffusional path length. Occlusion of the skin surface also increases skin temperature (~2 to 3°) resulting in increased molecular motion and skin permeation.

Hydrocarbon bases which occlude the skin to a degree will bring about an increase in drug penetration. However, this effect is trivial compared with the effects seen with a true occlusive skin wrap. Occlusive techniques are useful in some clinical situations requiring anti-inflammatory activity and occlusive wrappings are used most commonly with steroids. Since steroid activity can be enhanced so enormously by skin occlusion it is possible to depress adrenal function unknowingly. Early in the 1960s McKenzie demonstrated that penetration of steroid could be increased 100-fold by use of occlusion. The FDA requires the following label statement:

If extensive areas are treated or if the occlusive technique is used, the possibility exists of increased systemic absorption of the corticosteroid and suitable precautions should be taken.

Transdermal delivery systems, with their occlusive backing, can effect increased percutaneous absorption as a result of increased skin temperature and hydration.

Penetration Enhancers—This term has been used to describe substances that facilitate absorption through the skin. While most materials have a direct effect on the permeability of the skin, other so-called enhancers (eg polyols, such as glycerin and propylene glycol) appear to augment percutaneous absorption by increasing the thermodynamic activity of the penetrant, thereby increasing the effective escaping tendency and concentration gradient of the diffusing species. Penetration enhancers with a direct effect on skin permeability include solvents, surfactants and miscellaneous chemicals such as urea and *N,N*-diethyl-*m*-toluamide (Table III).[11] The mechanism of action of these enhancers is complex since these substances also may increase penetrant solubility. Nonetheless, the predominant effect of these enhancers on the stratum corneum is either to increase its degree of hydration or disrupt its lipoprotein matrix. In either case, the net result is a decrease in resistance to penetrant diffusion. (The formulator should note that the inclusion of a penetration enhancer in a topical formulation mandates additional testing and evaluation to ensure the absence of enhancer-related adverse effects.)

Foremost among the solvents which affect skin permeability is water. As noted above, water is a factor even for

Table III—Penetration Enhancers[a]

Solvents	
	Water
	Alcohols
	Methanol
	Ethanol
	2-propanol
	Alkyl methyl sulfoxides
	Dimethyl sulfoxide
	Decylmethyl sulfoxide
	Tetradecylmethyl sulfoxide
	Pyrrolidones
	2-Pyrrolidone
	N-Methyl-2-pyrrolidone
	N-(2-Hydroxyethyl)pyrrolidone
	Laurocapram
	Miscellaneous solvents
	Acetone
	Dimethyl acetamide
	Dimethyl formamide
	Tetrahydrofurfuryl alcohol
Amphiphiles	
	Anionic surfactants
	Cationic surfactants
	Amphoteric surfactants
	Nonionic surfactants
	Fatty acids and alcohols
Miscellaneous	
	Urea
	N,N-Diethyl-*m*-toluamide

[a] Adapted from Ref 11.

anhydrous transdermal delivery systems due to their occlusive nature. Due to its safety and efficacy, water has been described as the ultimate penetration enhancer. Other solvents include the classic enhancer, dimethyl sulfoxide (DMSO), which is of limited utility because of its potential ocular and dermal toxicity, its objectionable taste and odor (a consequence of its absorption and subsequent biotransformation) and the need for concentrations in excess of 70% to promote absorption. Analogs of DMSO such as decylmethyl sulfoxide are used currently in some topical formulations. In contrast with other solvents, laurocapram (1-dodecylazacycloheptan-2-one; Azone) has been shown to function effectively at low concentrations (≤5%). Furthermore, Azone's effect on skin permeability persists long after a single application due apparently to its prolonged retention within the stratum corneum.

Surfactants, long recognized for their ability to alter membrane structure and function, can have a substantial effect on skin permeability.[12] However, given the irritation potential of surfactants applied chronically, their utility as penetration enhancers is limited. Their effect on permeability may be complicated further by surfactant-monomer aggregation to form micelles and the concomitant solubilization of the permeant. As the impact of surfactants on skin permeability of a penetrant is problematic, the effect of their inclusion in a formulation should be evaluated using appropriate *in vitro* and *in vivo* studies.

Stratum Corneum Barrier Efficacy and Dermal Clearance—Even though *in vitro* studies of percutaneous transport may reflect the resistance of the skin to drug diffusion, there is no way such studies can characterize adequately the transfer of diffusing drug into the microvasculature of the dermis and its subsequent transfer into general circulation.

Christophers and Kligman[13] evaluated the dermal "clearance" of ^{22}Na from the midback skin of volunteers following the intradermal injection of ^{22}Na as normal saline solution. The dermal "clearances," expressed in terms of the half-life for disappearance of radioactivity, are plotted in Fig. 87-6.[13]

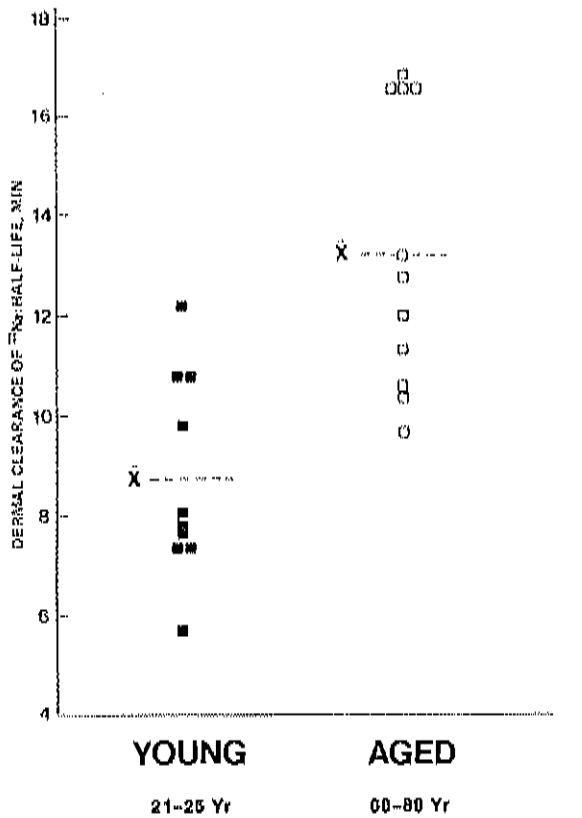


Fig 87-6. Dermal clearance of ^{22}Na in young and aged subjects after intradermal injection (data from Ref 13).

Similar results were obtained with disappearance of skin fluorescence after intradermal injection of sodium fluorescein. The data are indicative of markedly delayed dermal clearance in the aged. This may reflect, in part, a decrease in older subjects in dermal capillary loop density, a decrease in the rate and/or extent of dermal blood perfusion or an increase in resistance to transfer into the capillaries.

On the other hand, Christophers and Kligman[13] demonstrated increased *in vitro* skin permeation by sodium fluorescein in the stratum corneum excised from young and old subjects (Fig. 87-7[13]). Thus, the stratum corneum of older subjects may offer less resistance to the penetration of topically applied drugs.

Given the substantial intersubject variations that occur in diffusional resistance and in dermal clearance, it is not surprising that *in vivo* studies of percutaneous absorption often demonstrate marked differences in systemic availability of drugs. Furthermore, the tendency to employ normal, healthy, *young* adults in such studies may not provide data that is indicative of drug permeation through the skin of older subjects or patients. It would seem that more comprehensive studies of percutaneous absorption as a function of age are warranted.

Cutaneous Biotransformation—Catabolic enzyme activity in the viable epidermis is substantial. In fact, the viable epidermis is metabolically more active than the dermis. If the topically applied drug is subject to biotransformation during skin permeation, local and systemic bioavailability can be affected markedly. Enzymatic activity in the skin, or for that matter in systemic fluids and tissues, can be taken advantage of to facilitate percutaneous absorption. Sloan and Bodor,[14] for example, synthesized 7-acyloxymethyl derivatives of theophylline which diffuse through

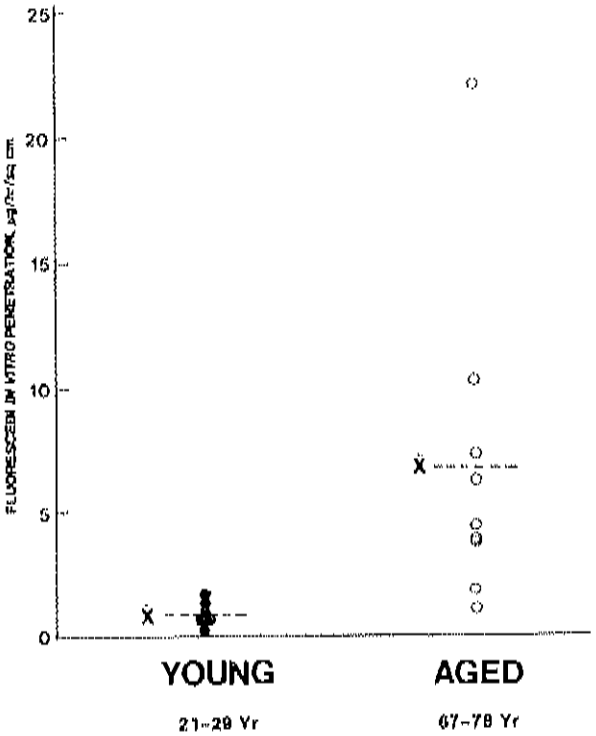


Fig 87-7. Flux of fluorescein through stratum corneum excised from young and aged subjects (data from Ref 13).

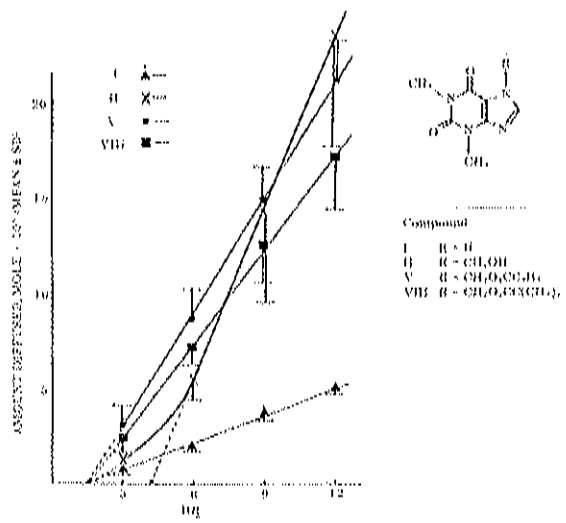


Fig 87-8. Diffusion of theophylline (I) and its derivatives through hairless mouse skin.[14]

the skin far more efficiently than theophylline itself (Fig 87-8[14]) but which are biotransformed rapidly to theophylline. Thus, theophylline delivery to systemic circulation can be enhanced substantially.

Further Considerations for Transdermal Drug Delivery

In order for a drug to qualify as a candidate for systemic delivery after topical application, it must satisfy requirements in addition to exhibiting good skin permeation. Successful candidates for transdermal drug delivery should be nonirritating and nonsensitizing to the skin. Since relatively little drug may reach systemic circulation over a relatively long time, drug candidates should be relatively potent drugs. In addition, the limitation to relatively potent drugs can ease problems of formulation since the amount of drug that can be incorporated in the formulation may be limited by physicochemical considerations such as solubility.

Iontophoretic Drug Delivery through the Skin[15,16]—For some poorly absorbed (ionic) compounds, parenteral administration appears to be the only viable option for regional or systemic delivery as chemical penetration enhancers (see Table III) often do not function well for these compounds. Given the increased risk of adverse reactions associated with the use of such enhancers, the increased evaluation of iontophoretic devices for the enhancement of topical drug delivery has been of great interest. Iontophoretic drug delivery implies the delivery of ionic drugs into the body by means of an electric current. While the stratum corneum forms the principal barrier to electrical conductivity—due, in part, to its lower water content—the skin also acts as a capacitor. Thus, biological tissues such as the skin provide for a reactive electrical circuit. Ionic transport through the skin in the presence of a uniform electric field can be described, in part, in accordance with the Nernst-Planck equation

$$J_i = -D\frac{dC}{dx} + \frac{DzeEC}{kT}$$

where J_i is the flux of ions across the membrane, C is the concentration of ions with valence, z, and electron charge, e, dC/dx is the concentration gradient, E is the electric field, k is Boltzmann's constant and T is the absolute temperature. Thus, the ionic flux is the sum of the fluxes that arise from the concentration gradient and the electric field. Given the complexity of the skin's composition, the thickness of the stratum corneum and the occurrence of electroosmotic effects, the Nernst-Planck equation is only a first approximation of the overall transdermal flux of a solute. Faraday's Law

$$\frac{Q}{t} = \frac{t_i I}{|z|F}$$

further characterizes the iontophoretic flux Q/t in terms of the current I (in amperes) and its duration t (in sec), the transference number parameter t_i and the Faraday constant F. Additional factors that influence the rate and extent of iontophoretic delivery through the skin include pH and ionic strength of the drug solution.

Although iontophoretic techniques have been shown to increase percutaneous absorption of ionizable or ionic drugs (including lidocaine, salicylates and peptides and proteins such as insulin) markedly, the clinical safety and efficacy of drug delivery systems employing iontophoretic technology have yet to be evaluated fully.

Ointments

Ointments are semisolid preparations intended for external application to the skin or mucous membranes; usually, but not always, they contain medicinal substances. The types of ointment bases used as vehicles for drugs are selected or designed for optimum delivery of the drugs and also to contribute emolliency or other quasi-medicinal qualities. Ointment properties vary, since they are designed for specific uses, ease of application or extent of application.

The official definition of ointment in its present form was introduced in the USP XV in 1955. The definition is broad and encompasses petrolatum, ie, oleaginous bases, emulsion bases—either water-in-oil (W/O) or oil-in-water (O/W)—and the so-called water-soluble bases.

In unofficial terms, oleaginous bases are described as ointments, but emulsion bases may be termed creams or lotions. Either of these containing large amounts of solids is termed a paste. All of these subclasses are defined officially as ointments.

Pharmaceutical authors have a penchant for defining "ideal" preparations eg, the ideal base, the ideal vehicle and so on. In practice, of course, there is no such thing. An individual cannot be all things to all people; neither can an ointment base be ideal for all drugs, all situations or all skins, for that matter. An ointment base functioning as a drug vehicle should be optimized for a specific drug and, insofar as possible, for specific disease states or skin conditions.

It is, of course, possible to define certain specific requirements for an ointment base to be used for extemporaneous compounding. Such a base should be nonirritating, easily removable, nonstaining, stable, non-pH-dependent and widely compatible with a variety of medicaments. When one adds the stipulation that the base must release the same variety of medicaments, the implausibility of such definitions becomes evident.

Classification and Properties of Ointment Bases

The USP recognizes four general classes of ointment bases, hereunder categorized into five classes for the purpose of indicating more definitively some differences in the principal properties of the bases.

Hydrocarbon Bases (Oleaginous)
Example: White Petrolatum
1. Emollient
2. Occlusive
3. Nonwater-washable
4. Hydrophobic
5. Greasy

Absorption Bases (Anhydrous)
Examples: Hydrophilic Petrolatum; Anhydrous Lanolin
1. Emollient
2. Occlusive
3. Absorb water
4. Anhydrous
5. Greasy

Emulsion Bases (W/O Type)
Examples: Lanolin, Cold Cream
1. Emollient
2. Occlusive
3. Contain water
4. Some absorb additional water
5. Greasy

Emulsion Bases (O/W Type)
Example: Hydrophilic Ointment
1. Water-washable
2. Nongreasy
3. Can be diluted with water
4. Nonocclusive

Water-Soluble Bases
Example: Polyethylene Glycol Ointment
1. Usually anhydrous
2. Water-soluble and washable
3. Nongreasy
4. Nonocclusive
5. Lipid-free

The selection of the optimum vehicle from the classification above may require compromises so often encountered in drug formulation. For example, stability or drug activity might be superior in a hydrocarbon base, however, acceptability is diminished because of the greasy nature of the base. The water-solubility of the polyethylene glycol bases may be attractive, but the glycol(s) may be irritating to traumatized tissue. Drug activity and percutaneous absorption may be superior when using a hydrocarbon base; however, it may be prudent to minimize percutaneous absorption by the use of a less occlusive base.

Ointment Bases

Hydrocarbon Bases

Hydrocarbon bases are usually petrolatum *per se* or petrolatum modified by waxes or liquid petrolatum to change viscosity characteristics. Liquid petrolatum gelled by the addition of a polyethylene resin also is considered a hydrocarbon ointment base, albeit one with unusual viscosity characteristics.

Hydrocarbon ointment bases are classified as oleaginous bases along with bases prepared from vegetable fixed oils or animal fats. Bases of this type include lard, benzoinated lard, olive oil, cottonseed oil and other oils. Such bases are emollient but generally require addition of antioxidants and other preservatives. They are now largely of historic interest.

Petrolatum USP is a tasteless, odorless, unctuous material with a melting range of 38 to 60°; its color ranges from amber to white (when decolorized). Petrolatum often is used externally, without modification or added medication, for its emollient qualities.

Petrolatum used as an ointment base has a high degree of compatibility with a variety of medicaments. Bases of this type are occlusive and nearly anhydrous and thus provide optimum stability for medicaments such as antibiotics. The wide melting range permits some latitude in vehicle selection and the USP permits addition of waxy materials as an aid in minimizing temperature effects.

Hydrocarbon bases, being occlusive, increase skin hydration by reducing the rate of loss of surface water. Bases of this kind may be used solely for such a skin-moisturizing effect, eg, white petroleum jelly as noted above. Skin hydration on the other hand may increase drug activity. Studies have indicated that steroids have increased activity, as measured by vasoconstrictor effects, when applied to the skin in a hydrocarbon vehicle. Stoughton consistently found the same steroid more active when applied in a petrolatum vehicle than when applied in a cream (ie, O/W emulsion) vehicle.

A gelled mineral oil vehicle represents a unique addition to this class of bases comprised of refined natural products. Liquid petrolatum may be gelled by addition of a polyethylene. When approximately 5% of low-density polyethylene is added, the mixture heated and then shock-cooled, a soft unctuous, colorless material resembling white petrolatum is produced. The mass maintains unchanged consistency over a wide temperature range. It neither hardens at low temperatures nor melts at reasonably high temperatures. Its useful working range is between −15° and 60°. Excessive heat, ie, above 90°, will destroy the gel structure.

On the basis of *in vitro* studies, drugs may be released faster from the gelled mineral oil vehicle than from conventional petrolatum. This quicker release has been attributed to easier migration of drug particulates through a vehicle which is essentially a liquid, compared with petrolatum.

Despite the advantages hydrocarbon or oleaginous vehicles provide in terms of stability and emolliency such bases have the considerable disadvantage of greasiness. The greasy or oily material may stain clothing and is difficult to remove. In terms of patient acceptance, hydrocarbon bases, ie, ointments, rank well below emulsion bases such as creams and lotions.

Absorption Bases

Absorption bases are hydrophilic, anhydrous materials or hydrous bases that have the ability to absorb additional

water. The former are anhydrous bases which absorb water to become W/O emulsions; the latter are W/O emulsions which have the ability to absorb additional water. The word absorption in this connotation refers only to the ability of the base to absorb water. Both types of base are exemplified by Anhydrous Lanolin and Lanolin. The former is converted to the latter by the addition of 30% water. The latter in turn will absorb additional amounts of water.

Hydrophilic Petrolatum USP is an anhydrous absorption base. The W/O emulsifying property is conferred by the inclusion of cholesterol. This composition is a modification of the original formulation which contained anhydrous lanolin. The lanolin was deleted because of reports of allergy; cholesterol was added. Inclusion of stearyl alcohol and wax add to the physical characteristics, particularly firmness and heat stability.

Hydrophilic Petrolatum USP

Cholesterol	30 g
Stearyl Alcohol	30 g
White Wax	80 g
White Petrolatum	860 g
To make	1000 g

Melt the stearyl alcohol and white wax together on a steam bath, then add the cholesterol and stir until it completely dissolves. Add the white petrolatum and mix. Remove from the bath, and stir until the mixture congeals.

Lanolin is a complex mixture of substances. Its ability to absorb water is probably a characteristic of the material rather than a single component. The chemistry of lanolin has been studied in detail. Such studies have resulted in the introduction of a large variety of lanolin derivatives and separated fractions. Available now are lanolin alcohols, dewaxed lanolins, acetylated lanolins, ethoxylated lanolins, hydrogenated lanolins, lanolin esters and other products. Most of these derivatives have been produced for specific purposes, such as improved emulsification characteristics or to reduce allergic reactivity.

The specific compounds responsible for lanolin allergy remain unknown; however, the greater portion of lanolin allergens reside in the wool wax alcohols fraction. Thus, fractional separation to obtain, for example, the so-called liquid lanolins substantially reduces the incidence of allergic reactions. Given the plethora of lanolin fractions, derivatives, modifications and levels of purity, it is quite possible, even likely, that lanolin-sensitive individuals can tolerate specific lanolin products.

Absorption bases, particularly the emulsion bases, impart excellent emolliency and a degree of occlusiveness on application. The anhydrous types can be used when the presence of water would cause stability problems with specific drug substances, eg, antibiotics. Absorption bases also are greasy when applied and are difficult to remove. Both of these properties are, however, less obvious than with hydrocarbon bases.

Commercially available absorption bases include Aquaphor (*Beiersdorf*) and Polysorb (*Fougera*). Nivea Cream (*Beiersdorf*) is a hydrated emollient base. Absorption bases, either hydrous or anhydrous, are seldom used as vehicles for commercial drug products. The W/O emulsion system is more difficult to deal with than the more conventional O/W systems and there is, of course, reduced patient acceptance because of greasiness.

Water-Removable Bases

Water-washable bases or emulsion bases, commonly referred to as creams, represent the most commonly used type of ointment base. By far the majority of commercial dermatologic drug products are formulated in an emulsion or cream base. Emulsion bases are washable and removed easily from skin or clothing. Emulsion bases can be diluted with water, although such additions are uncommon.

As a result of advances in synthetic cosmetic chemistry the formulator of an emulsion base can be faced with a bewildering variety of selections. Fortunately, the emulsion base can be subdivided into three component parts, designated as the oil phase, the emulsifier and the aqueous phase. The medicinal agent may be included in one of these phases or added to the formed emulsion.

The oil phase, sometimes called the internal phase, is typically made up of petrolatum and/or liquid petrolatum together with one or more of the higher-molecular-weight alcohols, such as cetyl or stearyl alcohol. Stearic acid may be included if the emulsion is to be based on a soap formed *in situ*, eg, triethanolamine stearate. A calculated excess of stearic acid in such a formulation will produce a pearlescent appearance in the finished product.

For drug-delivery vehicles, simplified systems are in order to minimize component interactions, either physical or chemical, and, of course, to minimize cost. Hydrophilic Ointment USP is a typical emulsion base. The composition is as follows:

Hydrophilic Ointment USP

Methylparaben	0.25 g
Propylparaben	0.15 g
Sodium Lauryl Sulfate	10 g
Propylene Glycol	120 g
Stearyl Alcohol	250 g
White Petrolatum	250 g
Purified Water	370 g
To make about	1000 g

Melt the stearyl alcohol and the white petrolatum on a steam bath, and warm to about 75°. Add the other ingredients, previously dissolved in the water and warmed to 75°, and stir the mixture until it congeals.

Stearyl alcohol and petrolatum comprise an oil phase with the proper smoothness and comfort for the skin. Stearyl alcohol also serves as an adjuvant emulsifier. Petrolatum in the oil phase also contributes to the water-holding ability of the overall formulation.

A glance at the cosmetic literature and such volumes as the Cosmetic, Toiletry and Fragrance Association's *Cosmetic Ingredient Dictionary* impresses one with the enormous number and variety of emulsion-base components, particularly oil-phase components. Many of these substances impart subtle but distinct characteristics to cosmetic emulsion systems. While desirable, many of these characteristics are not really necessary in drug dosage forms and delivery systems.

The aqueous phase of an emulsion base usually, but not always, exceeds the oil phase in volume. The aqueous phase contains the preservative materials, the emulsifier or a part of the emulsifier system and humectant. The last is usually glycerin, propylene glycol or a polyethylene glycol. The humectant normally is included to minimize water loss in the finished composition. Humectants also add to overall physical product acceptability.

The aqueous phase contains the preservative(s) which are included to control microbial growth. Preservatives in emulsion bases usually include one or more of the following: methylparaben and propylparaben, benzyl alcohol, sorbic acid or quaternary ammonium compounds. Propylene glycol in sufficient concentration also can function as a preservative. The general subject of preservatives and preservation is discussed elsewhere in this chapter.

The aqueous phase also contains the water-soluble components of the emulsion system, together with any additional stabilizers, antioxidants, buffers, etc that may be neces-

sary for stability, pH control or other considerations associated with aqueous systems.

The emulsifier or emulsifier system in a cream formulation is a major consideration. The emulsifier may be nonionic, anionic, cationic or amphoteric.

Anionic Emulsifiers—Sodium lauryl sulfate, the emulsifier in Hydrophilic Ointment USP, is typical of this class. The active portion of the emulsifier is the anion (lauryl sulfate ion). Similar anionic emulsifiers include soaps such as triethanolamine stearate. Soaps, of course, are alkaline and, hence, incompatible with acids.

Sodium lauryl sulfate and other anionic surfactants of its type are more acid-stable and permit adjustment of the emulsion pH to the desirable acid range of 4.5 to 6.5. As anionic emulsifiers are incompatible with cations, the overall product composition must be kept in mind.

Depending on the chemical type and concentration, anionic surfactants may be irritating in certain situations. It has been reported that percutaneous absorption of certain drugs, notably steroids, may be enhanced by the use of anionic compounds such as sodium lauryl sulfate.

Cationic Emulsifiers—Cationic compounds are highly surface-active but are used infrequently as emulsifiers. The cation portion of the molecule is generally a quaternary ammonium salt including (usually) a fatty acid derivative, eg, dilauryldimethylammonium chloride. Cationics may be irritating to the skin and eyes, and they have a considerable range of incompatibilities, including anionic materials.

Nonionic Emulsifiers—Nonionic emulsifiers show no tendency to ionize in solution. This advantage results in excellent pH and electrolyte compatibility in such emulsions. Nonionic emulsifiers range from lipophilic to hydrophilic. The usual emulsifier system may include both a lipophilic and hydrophilic member to produce a so-called hydrophilic-lipophilic balance (or HLB).

Many nonionic surfactants are the result of condensation of ethylene oxide groups with a long chain hydrophobic compound. The hydrophilic characteristics of the condensation product are controlled by the number of (usually) oxyethylene groups (OCH_2CH_2). Examples of nonionic surfactants are given in Table IV.[17]

Emulsions containing nonionic emulsifiers usually are prepared by dissolving or dispersing the lipophilic component in the oil phase and the hydrophilic component in the aqueous phase. The two phases then are heated separately and combined as described on page 1534. The nonionic emulsifier content of an emulsion may total as much as 10% of the total weight or volume. Emulsions based on nonionic emulsifiers are generally low in irritation potential, stable and have excellent compatibility characteristics.

Soaps and detergents, ie, emulsifiers, have, overall, a damaging effect on the skin. Both anionic and cationic surfactants can cause damage to the stratum corneum in direct proportion to concentration and duration of contact. Nonionic surfactants appear to have much less effect on the stratum corneum.

After the proper selection of ingredients the emulsion base is formed by heat and agitation. The oil phase is melted and heated to 75° in a container equipped with a variable-speed agitator. The aqueous phase with the emulsifier added is placed in a second container, components are dissolved and the whole heated to 75° or slightly in excess. The aqueous phase then is added slowly with continuous stirring to the oil phase. The first addition should be carried out slowly but continuously with thorough but careful agitation, ie, the emulsion should not be agitated at a rate that incorporates excess air. Progressively slower stirring should be continued during addition of the aqueous phase and until the temperature reaches about 30°. Medicinal agents usually are added after the emulsion has formed and much of the aqueous phase has been added. Drug substances frequently are added as dispersed concentrates in aqueous suspension. Colors and dyes are similarly added as concentrates. Colors sometimes are employed to distinguish different concentrations of the same drug product. Fragrances, if any, are added after the formed emulsion has cooled to about 35°.

Table IV—Nonionic Emulsifiers[17]

Type	Examples
Polyoxyethylene fatty alcohol ethers	Polyoxyethylene lauryl alcohol
Polyoxypropylene fatty alcohol ethers	Propoxylated oleyl alcohol
Polyoxyethylene fatty acid esters	Polyoxyethylene stearate
Polyoxyethylene sorbitan fatty acid esters	Polyoxyethylene sorbitan monostearate
Sorbitan fatty acid esters	Sorbitan monostearate
Polyoxyethylene glycol fatty acid esters	Polyoxyethylene glycol monostearate
Polyol fatty acid esters	Glyceryl monostearate Propylene glycol monostearate
Ethoxylated lanolin derivatives	Ethoxylated lanolins Ethoxylated cholesterol

Water-Soluble Bases

Soluble ointment bases, as the name implies, are made up of soluble components, or may include gelled aqueous solutions. The latter often are referred to as gels, and in recent years have been formulated specifically to maximize drug availability.

Major components, and in some instances the only components, of water-soluble bases are the polyethylene glycols. These are liquids or waxy solids identified by numbers which are an approximate indication of molecular weight. Polyethylene glycol 400 is a liquid superficially similar to propylene glycol, while polyethylene glycol 4000 is a waxy solid.

Polyethylene glycols have the general chemical formula

$$HOCH_2(CH_2OCH_2)_nCH_2OH$$

They are nonvolatile, water-soluble or water-miscible compounds and chemically inert, varying in molecular weight from several hundred to several thousand. Patch tests have shown that these compounds are innocuous and continuous use has confirmed their lack of irritation.

Polyethylene glycols of interest as vehicles include the 1500, 1600, 4000 and 6000 products, ranging from soft, waxy solids (polyethylene glycol 1500 is similar to petrolatum) to hard waxes. Polyethylene glycol 6000 is a hard wax-like material melting at 58 to 62°; it is nonhygroscopic.

Polyethylene glycols, particularly 1500, can be used as a vehicle *per se;* however, better results often are obtained by using blends of high- and low-molecular-weight glycols, as in Polyethylene Glycol Ointment NF.

Polyethylene Glycol Ointment NF

Polyethylene Glycol 3350	400 g
Polyethylene Glycol 400	600 g

Heat the two ingredients on a water bath to 65°. Allow to cool and stir until congealed. If a firmer preparation is desired, replace up to 100 g of the polyethylene glycol 400 with an equal amount of polyethylene glycol 3350.

Note—If 6–25% of an aqueous solution is to be incorporated in polyethylene glycol ointment, replace 50 g of the polyethylene glycol 3350 with an equal amount of stearyl alcohol.

The water-solubility of polyethylene glycol vehicles does not insure availability of drugs contained in the vehicle. As hydrated stratum corneum is an important factor in drug penetration, the use of polyethylene glycol vehicles which

are anhydrous and nonocclusive actually may hinder percutaneous absorption due to dehydration of the stratum corneum.

Aqueous gel vehicles containing water, propylene and/or polyethylene glycol, and gelled with a carbomer or a cellulose derivative, also are classed as water-soluble bases. Bases of this kind, sometimes referred to as gels, may be formulated to optimize delivery of a drug, particularly steroids. In such a preparation propylene glycol is used as a steroid solvent as well as an antimicrobial or preservative.

Gelling agents used in these preparations may be nonionic or anionic. Nonionics include cellulose derivatives, such as methylcellulose or hydroxypropyl methylcellulose. These derivatives form gels when dissolved in water but also exhibit the characteristic of reverse solubility. The celluloses are wetted, ie, dispersed in hot water, and then cooled to effect solution. Sodium carboxymethylcellulose is an ionic form of cellulose gelling agent. It is conventionally soluble, and not heat-insoluble.

Carbopol 934 is a white, fluffy, powdered polymeric acid, dispersible but insoluble in water. When the acid dispersion is neutralized with a base a clear, stable gel is formed. Carbopol 934 is physiologically inert and is not a primary irritant or sensitizer.

Another gelling agent is colloidal magnesium aluminum silicate (*Veegum*). It is an inorganic emulsifier and suspending agent, as well as a gelling agent. Veegum dispersions are compatible with alcohols (20 to 30%), acetone and glycols. It frequently is employed as a gel stabilizer, rather than as the sole gelling agent.

Sodium alginate and the propylene glycol ester of alginic acid (*Kelcoloid*) also are satisfactory gelling agents. Sodium alginate is a hydrophilic colloid that functions satisfactorily between pH 4.5 and 10; addition of calcium ions will gel fluid solutions of sodium alginate.

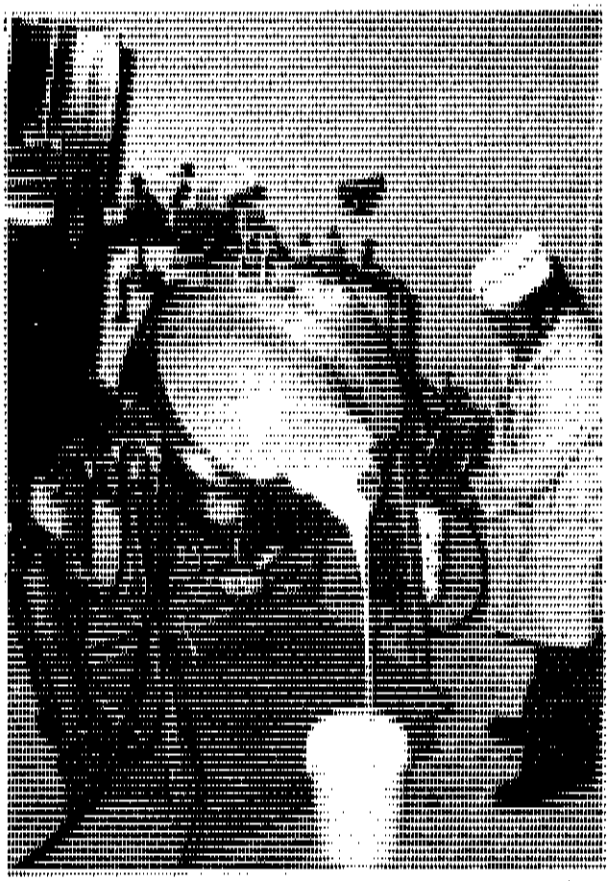


Fig 87-9. Pilot scale ointment manufacture (courtesy, Alcon).

Preparation

Ointment preparation or manufacture depends on the type of vehicle and the quantity to be prepared. The objective is the same, ie, to disperse uniformly throughout the vehicle a finely subdivided or dissolved drug substance(s). Normally, the drug materials are in finely powdered form before being dispersed in the vehicle.

Incorporation by Levigation

The preparation of small quantities of ointment by the pharmacist, ie, one to several ounces, can be accomplished by using a spatula and an ointment tile (either porcelain or glass). The finely powdered drug material is levigated thoroughly with a small quantity of the base to form a concentrate. The concentrate then is diluted geometrically with the remainder of the base. Such a procedure is useful particularly with petrolatum or oleaginous bases.

If the drug substance is water-soluble it can be dissolved in water and the resulting solution incorporated into the vehicle using a small quantity of lanolin if the base is oleaginous. Generally speaking, an amount of anhydrous lanolin equal in volume to the amount of water used will suffice.

When ointments are made by incorporation in quantities too large to be handled with a tile and spatula, mechanical mixers are used. Hobart mixers, pony mixers and others of the type usually are used for this purpose. The drug substance in finely divided form usually is added slowly or sifted into the vehicle contained in the rotating mixer. When the ointment is uniform, the finished product may be processed through a roller mill to assure complete dispersion and reduce any aggregates.

This procedure may be modified by preparing and milling

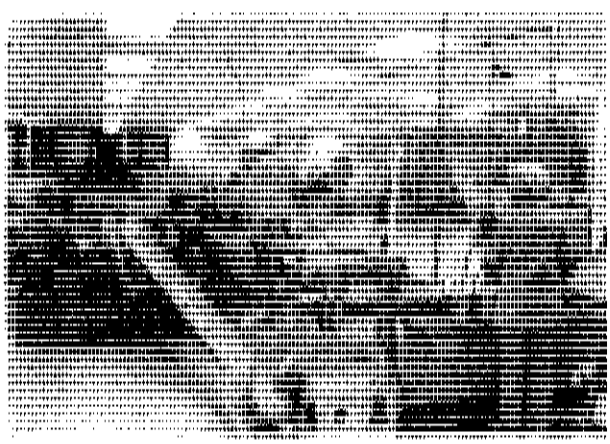


Fig 87-10. Ointment manufacture and packaging (courtesy, Owen Laboratories).

a concentrate of the drug in a portion of the base. The concentrate then is dispersed in the balance of the vehicle, using a mixer of appropriate size. Occasionally, the base may be melted for easier handling and dispersing. In such cases the drug is dispersed and the base slowly cooled using continuous agitation to maintain dispersion.

Emulsion Products

Medicated creams and lotions are prepared by means of a two-phase heat system. The oil-phase ingredients are combined in a jacketed tank and heated to about 75°. At this

temperature the oil-phase ingredients are liquefied and uniform. In a separate tank the aqueous-phase ingredients, including the emulsifier, are heated together to slightly above 75°. The aqueous phase then is added to the oil phase, slowly and with constant agitation. When the emulsion is formed the mixture is allowed to cool, maintaining slow agitation.

At this stage in the process the medicinal ingredients usually are added as a concentrated slurry, which usually has been milled to reduce any particle aggregates. Volatile or aromatic materials generally are added when the finished emulsion has cooled to about 35°. At this point additional water may be added to compensate for any evaporative losses occurring during exposure and transfer at the higher temperatures of emulsion formation.

While the product remains in the tank in bulk, quality-control procedures are performed, ie, for pH, active ingredients, etc. If control results are satisfactory the product is filled into the appropriate containers.

Preservatives in Ointment Bases

Antimicrobial preservative substances are included in ointment formulations to maintain the potency and integrity of product forms and to protect the health and safety of the consumer. The USP addresses this subject in its monograph on Microbiological Attributes of Non-Sterile Pharmaceutical Products. The significance of microorganisms in nonsterile products should be evaluated in terms of the use of the product, the nature of the product and the potential hazard to the user. The USP suggests that products applied topically should be free from *P aeruginosa* and *S aureus*.

The attributes of an ideal preservative system have been defined by various authors as

1. Effective at relatively low concentrations against a broad spectrum or variety of microorganisms which could cause disease or product deterioration.
2. Soluble in the required concentration.
3. Nontoxic and nonsensitizing at in-use concentrations.
4. Compatible with ingredients of the formulation and package components.
5. Free from objectionable odors and colors.
6. Stable over a wide spectrum of conditions.
7. Inexpensive.

No preservative or preservative system meets these ideal criteria. In fact, preservative substances once considered most acceptable, if not ideal, now have been questioned. Methylparaben and propylparaben, second and third only to water in frequency of use in cosmetic formulations, have been associated with allergic reactions.

Use of parabens as preservatives in topical products began nearly a half-century ago. Animal testing indicated that they virtually are nontoxic and the compounds, usually in combination, became nearly ubiquitous as preservatives in dermatologic and cosmetic products. In 1968 Schorr was among the first in this country to express concern about contact sensitization to parabens. Other investigators have voiced similar concerns.

Topical parabens do not appear to constitute a significant hazard to the public based on their low index of sensitization and low overall toxicity.

Alternative preservation substances available for use in ointment bases, together with comments on possible limitations, are given in Table V.[18] It is probably sensible to note that, with few exceptions, most of these compounds—in contrast to the parabens—do not have a half-century history of use nor have had extensive patch-testing experiments carried out.

Following selection of preservative candidates and preparation of product prototypes, the efficacy of the preservative system must be evaluated. A variety of methods to accomplish this have been proposed. The organism challenge procedure is currently the most acceptable. In this procedure the test-product formulation is inoculated with specific levels and types of microorganisms. Preservative efficacy is evaluated on the basis of the number of organisms killed or whose growth is inhibited as determined during a specific sampling schedule. Critical to the organism challenge procedure are the selection of challenge microorganisms, the level of organisms in the inoculum, the sampling schedule and data interpretation.

Variations of the organism challenge procedure have usually centered around the selection of organisms, the challenge schedule, use of a rechallenge and standards of effectiveness, ie, cidal activity required rather than static or inhibitory activity.

Table VI gives the challenge organisms and other criteria used in several preservative challenge procedures.

In addition to efficacy in terms of antimicrobial effects, the preservative system must be assessed in terms of chemical and physical stability as a function of time. This often is done using antimicrobial measurements in addition to chemical analysis.

Table V—Topical Preservatives: Benefits and Risks[18]

Preservatives	Limitations relative to use in cosmetic/dermatological formulations
Quaternary ammonium compounds	a) inactivated by numerous ingredients including anionics, nonionics and proteins
Organic mercurial compounds	a) potentially toxic and may sensitize the skin b) limited use in formulations used near or in the eye
Formaldehyde	a) volatile compound with an objectionable odor b) irritating to the skin c) high chemical reactivity
Halogenated phenols hexachlorophene, *p*-chloro-*m*-cresol (PCMC) *p*-chloro-*m*-xylenol (PCMX) dichloro-*m*-xylenol (DCMX)	a) objectionable odor b) often inactivated by nonionics, anionics or proteins c) limited gram-negative antibacterial activity
Sorbic acid potassium sorbate	a) pH-dependent (can be used only in formulations below the pH of 6.5 to 7.0) b) higher concentrations are oxidized by sunlight resulting in product discoloration c) limited antibacterial activity
Benzoic acid sodium benzoate	a) pH-dependent (limited to use in formulations with pH of 5.5 or less) b) replaced by newer antimicrobials because of its limited antimicrobial activity

Table VI—Preservative Effectiveness Test Procedures

	USP XX	CTFA	FDA
Challenge microorganisms	*S aureus* *E coli* *P aeruginosa* *C albicans* *A niger*	*S aureus* *E coli* *P aeruginosa* *C albicans* *A niger* *P luteum* *B subtilis*	*S aureus* *E coli* *P aeruginosa* *P putida* *P multivorans* *Klebsiella* *S marcescens* *C albicans* *A niger*
Inoculum level	1×10^5–1×10^6 Cells/mL or gm	1×10^6 Cells/mL or gm	0.8–1.2×10^6 Cells/mL or gm rechallenge 1–2.0×10^5 vegetative cells
Sampling schedule	0, 7, 14, 21, 28 days	0, 1–2, 7, 14, 28 days	weekly intervals
Standards	Bacteria < 0.1% survival by 14th day. Yeast & molds at or below initial concentration during first 14 days. No increase in organism counts for remainder of 28-day survival	Based on intended use	Vegetative cells < 0.01% survival in 28 days *C albicans* < 1% survival *A niger* < 1% survival Rechallenge 0.1% survival in 28 days

Safety, Safety Testing and Toxicity

Safety is defined as the condition of being safe from undergoing (or causing) injury. Safety is not absolute but must be taken in the context of conditions of use. Toxicity refers to a specific substance or product and the adverse effect on a system caused by such a substance or product acting for a given period of time at a specific dose level.

Ointment bases may cause irritant or allergic reactions. Allergic reactions are usually to a specific base component. Irritant reactions are more frequent and more important, hence a number of test procedures have been devised to test for irritancy levels, both in the animal and in man. The consequences of species differences and specificity must be included in the evaluation of animal-test results.

Probably the most common irritancy measure is the Draize dermal irritation test in rabbits. In this procedure the test material is applied repeatedly to the clipped skin on the rabbit's back. The test material may be compared with one or more control materials.

End-points are dermal erythema and/or edema. By assigning numerical scores for erythema and edema, mathematical and statistical treatment of results is possible.

In the human, a variety of test procedures are used to measure irritancy, sensitization potential and phototoxicity. Among the most common are the following:

21-Day Cumulative Irritation Study

In this test the test compound is applied daily to the same site on the back or volar forearm. Test materials are applied under occlusive tape and scores are read daily. The test application and scoring is repeated daily for 21 days or until irritation produces a predetermined maximum score. Typical erythema scores are

0 = no visible reaction
1 = mild erythema
2 = intense erythema
3 = intense erythema with edema
4 = intense erythema with edema and vesicular erosion.

Usually, 24 subjects are used in this test. Fewer subjects and a shorter application time in days are variants of the test.

Draize-Shelanski Repeat-Insult Patch Test

This test is designed to measure the potential to cause sensitization. The test also provides a measure of irritancy potential. In the usual procedure the test material or a suitable dilution is applied under occlusion to the same site, for 10 alternate-day 24-hr periods. Following a 7-day rest period the test material is applied again to a fresh site for 24 hr. The challenge sites are read on removal of the patch and again 24 hr later. The 0–4 erythema scale is used. A test panel of 100 individuals is common.

Kligman Maximization Test

This test is used to detect the contact sensitizing potential of a product or material. The test material is applied under occlusion to the same site for 48-hr periods. Prior to each exposure the site may be pretreated with a solution of sodium lauryl sulfate under occlusion. Following a 10-day interval the test material again is applied to a different site for 48 hr under occlusion. The challenge site may be treated briefly with a sodium lauryl sulfate solution.

The Maximization test is of shorter duration and makes use of fewer test subjects than the Draize Shelanski test. The use of sodium lauryl sulfate as a pretreatment increases the ability to detect weaker allergens.

These test methods are adequate to detect even weak irritants and weak contact sensitizers. Positive results, however, automatically do not disqualify the use of a substance as unsafe. The actual risk of use depends on concentration, period of use and skin condition. Benzoyl peroxide in tests such as the Draize Shelanski and Maximization is a potent sensitizer, yet the incidence of sensitization among acne patients is low.

Packaging and Labeling

Ointments usually are packaged in ointment jars or in metal or plastic tubes of a convenient size. Ointment jars are available in one-half to 16-oz sizes; tubes from 3.5-g capacity (often ophthalmic) to 4-oz and on occasion greater capacities.

Ointment Jars—Straight-sided screw cap jars of glass or plastic are available. Clear, amber or opaque glass containers are used, as well as white, opaque, plastic, usually high-density polyethylene, jars. Metal or composition plastic tops are available, with a variety of inner liners to assure a dust- and airtight closure. Liners are usually paper or plastic laminates or discs glued or otherwise fitted to the closure.

Ointment jars are filled mechanically to somewhat less than capacity to minimize contact between the ointment and the cap or cap-liner. Ointment jars hand-filled by the pharmacist also should be finished to avoid contact between the

ointment and cap. This can be accomplished quite readily by skillful use of a flexible spatula. The spatula is forced across the ointment jar while depressed slightly into the ointment. The result is a conical depression that is esthetically acceptable. Much of the same result can be accomplished by depressing the spatula into the center of the filled jar and gradually rotating the jar against the stationary spatula. Small points perhaps, but time well spent to avoid having part of the ointment-jar contents removed inadvertently by the cap when the patient opens the jar.

Ointment Tubes—Ointment tubes made of tin or aluminum, or of an increasing variety of plastic materials, are available. The latter are normally polyethylene, polypropylene or other flexible, heat-sealable plastics. Ointment tubes have obvious advantages over jars; the use of fingers is minimized, as is dust and air contact, and light exposure.

Depending on the expected shelf-life, a number of factors should be considered in selecting an ointment tube. Metal contact and the possibility of metal-ion catalyzed instability must be considered. Conversely, plastic tubes may become stained or discolored by migration of colored materials into the plastic sidewalls of the tube; coal tar in ointment form may cause such discoloration. Tube interactions involving either metal or plastic can be minimized by internal coatings. Such coatings usually are epoxy films that become the primary product contact.

The suitability of ointment containers, either jars or tubes, should be verified by adequate testing prior to use. Compatibility and physical and chemical stability should be established by proper tests before final selection of a jar or tube.

Ointments prepared on prescription can be conveniently filled into a metal ointment tube using the following procedure.

Select an ointment tube of the proper size and remove any lint or dust. Transfer the ointment to a piece of paper of suitable size (use glassine or strong paper). Roll the paper and ointment into a cylinder shape of a diameter slightly less than that of the ointment tube. Insert the rolled paper-ointment cylinder into the ointment tube. The length of the paper cylinder should exceed the tube length. Remove the ointment tube cap and, using a spatula, compress the paper cylinder and tube. Continue compressing the ointment and tube until the ointment appears in the neck orifice of the open tube. Replace the cap. Using the spatula side as a knife-edge, compress the ointment tube and paper cylinder a reasonable distance from the end of the tube. Holding the spatula firmly in place, draw out the paper cylinder, leaving the ointment within the tube.

The ointment tube selected should be of adequate capacity. After compressing the ointment and paper cylinder into the tube, constrict the tube for cylinder removal at a distance from the end of the tube that will allow at least a double foldover to seal the tube. The fold dimensions are inexact, however, the individual folds on a 1-oz tube are approximately $\frac{1}{16}$ to $\frac{3}{16}$ in. Ointment tube sealing folds easily can be made by folding the tube over on itself using a spatula blade to flatten the tube and serve as a folding point. Ointment tube clips can be fixed over the tube ends and clamped in place using pliers or a small vise. The sole purpose of folding and clamping is to prevent leakage when routine use pressure is applied to the tube.

On a larger scale, ointment-tube filling is accomplished using automatic equipment which air-cleans the tubes, fills, folds and crimps the end in one continuous operation. Some equipment will stamp an expiration date onto the crimped surface. In larger-scale manufacturing operations plastic tubes are used with increasing frequency. From a filling standpoint plastic tubes are handled much like metal tubes. The final step, however, is a heat seal with no end foldover.

Labeling Ointment Tubes—Attaching labels to ointment tubes is a minor difficulty compounded by the increasing unsightliness characteristic of many ointment tubes during use. The label increasingly can become obliterated, difficult to read and, frequently, lost. As a general rule the label should be attached to itself, ie, it should completely encircle the tube. It should be attached to the tube, affixed close to the neck end.

Given the usual handling of ointment tubes by the patient, it is good practice to dispense the tube in a vial or hinged pasteboard box of convenient size. The outer container serves to hold and protect the ointment tube as well as to carry the label. The ointment tube is marked with a container prescription number so that both tube and container are identified.

On a manufacturing scale tubes are labeled in a variety of ways. Paper labels may be used, labeling may be silk-screened onto plastic surfaces; expiration dates and code lot numbers may be stamped on as a part of the tube-crimping procedure.

Suppositories

Suppositories are solid dosage forms of various weights and shapes, usually medicated, for insertion into the rectum, vagina or the urethra. After insertion, suppositories soften, melt or dissolve in the cavity fluids.

The use of suppositories dates from the distant past, this dosage form being referred to in writings of the early Egyptians, Greeks and Romans. Suppositories are suited particularly for administration of drugs to the very young and the very old, a notion first recorded by Hippocrates. Despite the antiquity of this dosage form, little was known about drug absorption or drug activity via suppository administration until recent years.

Types

Rectal Suppositories—The USP describes rectal suppositories for adults as tapered at one or both ends and usually weighing about 2 g each. Infant rectal suppositories usually weigh about one-half that of adult suppositories. Drugs having systemic effects, such as sedatives, tranquilizers and analgesics, are administered by rectal suppository; however, the largest single-use category is probably that of hemorrhoid remedies dispensed over-the-counter. The 2-g weight for adult rectal suppositories is based on use of cocoa butter as the base; when other bases are used the weights may be greater or less than 2 grams.

Vaginal Suppositories—The USP describes vaginal suppositories as usually globular or oviform and weighing about 5 g each. Vaginal medications are available in a variety of physical forms, eg, creams, gels or liquids, which depart from the classical concept of suppositories. Vaginal tablets, however, do meet the definition, and represent convenience both of administration and manufacture.

Urethral Suppositories—Urethral suppositories—sometimes referred to as bougies—are not described specifically in the USP, either by weight or dimension. Traditional values, based on use of cocoa butter as base, are as follows for these cylindrical dosage forms: diameter, 5 mm; length, 50 mm female, 125 mm male; weight, 2 g female, 4 g male. Urethral suppositories are an unusual dosage form and seldom are encountered.

Rectal Absorption

Drug absorption for systemic activity generally is limited to rectal administration. As noted previously, the bioavailability of rectally administered drugs is a relatively recent concern. Rectally instilled preparations, whether supposi-

tories, foams or solutions (enemas), tend to be confined to the rectum and sigmoid colon if the volume is less than about 50 mL. Foams tend to dissipate or spread to a lesser extent than solutions, particularly large-volume solutions (~100 to 200 mL). Though large-volume fluid formulations—solutions or enemas—may allow drug to reach the ascending colon, substantial intra- and intersubject variation is evident.[19] Literature information indicates that rectal drug absorption from suppositories can be erratic and may be substantially different from absorption following oral administration. With only a few recent exceptions, suppository studies are based on either *in vivo* or *in vitro* data with few attempts to correlate *in vitro* results with *in vivo* studies.

Major factors affecting the absorption of drugs from suppositories administered rectally are the following: anorectal physiology, suppository vehicle, absorption site pH, drug pK_a, degree of ionization and lipid solubility.

Anorectal Physiology—The rectum is about 150 mm in length, terminating in the anal opening. In the absence of fecal matter the rectum contains a small amount of fluid of low buffering capacity. Fluid pH is said to be about 7.2; because of the low buffer capacity pH will vary with the pH of the drug product or drug dissolved in it. The rectal epithelium is lipoidal in character. The lower, middle and upper hemorrhoidal veins surround the rectum. Only the upper vein conveys blood into the portal system, thus drugs absorbed into the lower and middle hemorrhoidal veins will bypass the liver. Absorption and distribution of a drug therefore is modified by its position in the rectum, in the sense that at least a portion of the drug absorbed from the rectum may pass directly into the inferior vena cava, bypassing the liver.

Spreading characteristics of rectal formulations may be affected considerably by intraluminal rectal pressure—due, in part, to the weight of abdominal organs and to respiratory activity—and by periodic contractile activity of the rectal wall.[20]

Suppository Vehicle—The ideal suppository base should meet the following general specifications:

1. The base is nontoxic and nonirritating to mucous membranes.
2. The base is compatible with a variety of drugs.
3. The base melts or dissolves in rectal fluids.
4. The base should be stable on storage; it should not bind or otherwise interfere with release and absorption of drug substances.

Absorption Factors—Prior to absorption the administered drug must be in solution. Solution, therefore, must be preceded by dissolution or melting of the vehicle. Dissolution is followed by partitioning or diffusion of the drug into the rectal fluid.

Rectal suppository bases can be classified broadly into two types. The traditional cocoa butter vehicle is immiscible with aqueous tissue fluids but melts at body temperature. Water-soluble vehicles also have been used. Typical of this class is the polyethylene glycol vehicle. Drug absorption from such dissimilar bases can differ substantially. Lowenthal and Borzelleca[21] investigated the absorption of salicylic acid and sodium salicylate administered to dogs. The drugs were formulated in a cocoa butter base and in a base comprised of polyethylene glycol, synthetic glycerides and a surfactant. Absorption of salicylic acid and sodium salicylate was about equal from the cocoa butter base; however, salicylic acid gave higher plasma levels than sodium salicylate when the glycol base was used.

Parrott[22] compared the absorption of salicylates after rectal and oral administration. Using urinary excretion data both aspirin and sodium salicylate were found to be equally bioavailable orally or rectally. Aspirin was released more rapidly from water-miscible suppositories than from the oily type. Conversely, sodium salicylate was released more rapidly from a cocoa butter vehicle.

Based on available data the bioavailability of a drug from a suppository dosage form is dependent on the physicochemical properties of the drug as well as the composition of the base. The drug-dissolution rate and, where appropriate, the partition coefficient between lipid and aqueous phase should be known.

For suppository formulation the relative solubility of the drug in the vehicle is a convenient comparison measure. Lipid-soluble drugs present in low concentration in a cocoa butter base will have little tendency to diffuse into rectal fluids. Drugs that are only slightly soluble in the lipid base will partition readily into the rectal fluid. The partition coefficient between suppository base and rectal fluid thus becomes a useful measure. In water-soluble bases and assuming rapid dissolution, the rate-limiting step in absorption would be transport of the drug through the rectal mucosa.

A wide variety of substances have been investigated for their ability to enhance rectal permeability to drugs. Agents such as EDTA have been used to chelate Ca^{2+} and Mg^{2+} in the vicinity of paracellular tight junctions and, thus, alter epithelial permeability. Other promoters of rectal absorption (eg, bile salts and nonsteroidal anti-inflammatory agents, including aspirin, salicylic acid and diclofenac) appear to exert their influence by affecting water influx and efflux rates across the rectal mucosa. Surfactants not only may modify membrane permeability but also enhance wetting or spreading of the base and dissolution of the drug. In any event, it should be evident that, whatever the mechanism, enhancing the *rectal* absorption of drugs—especially those which undergo presystemic elimination—could result in substantially reduced dosage requirements and decreased risk of adverse reactions.

Clearly, the bioavailability of a drug administered rectally depends on the nature of the drug and the composition of the vehicle or base. The physical properties of the drug can be modified to a degree, as can the characteristics of the base selected as the delivery system. Preformulation evaluations of physicochemical properties must then be confirmed by *in vivo* studies in animals and ultimately in the primary primate, man.

***In Vivo* Rectal Absorption Studies**—Dogs are probably the animal of choice in evaluating rectal drug availability. (The pig is a closer physiological match, but size and manageability argue in favor of the dog.) Blood and urine samples can be obtained from the dog and rectal retention can be accomplished with facility. Smaller animals have been used; rabbits, rats and even mice have been employed, but dosing and sampling become progressively more difficult.

Human subjects provide the ultimate measure of drug bioavailability. Subjects are selected on the basis of age, weight and medical history. Subjects usually are required to fast overnight and evacuate the bowel prior to initiation of the study. Fluid volume and food intake usually are standardized in studies of this kind.

Given the difficulty of standardizing pharmacological end-points the usual measure of rectal drug bioavailability is the concentration of the drug in blood and or urine as a function of time. A control group using oral drug administration provides a convenient means of comparing oral and rectal drug availability. Such a comparison is meaningful particularly in view of uncertainties and conflicts encountered in the literature. While there is general agreement about drug absorption from the rectum there is less agreement on dosage adequacy and the relationship between oral and rectal dosage. This state of affairs argues in favor of adequate studies to establish proper dosage and verify bioavailability.

Vaginal Absorption

Passive drug absorption via the vaginal mucosa, as with other mucosal tissues, is influenced by absorption site physiology, absorption site pH and the solubility and partitioning characteristics of the drug. The vaginal epithelial surface usually is covered with an aqueous film—emanating from cervical secretions—whose volume, pH and composition vary with age, stage of the menstrual cycle and location. Postmenarche, a vaginal pH gradient is evident with the lowest values (pH ~4.0) near the anterior fornix and the highest (pH ~5.0) near the cervix.[23] Following intravaginal administration, some drug absorption from the intact vaginal mucosa is likely, even when the drug is employed for a local effect. In fact, extensive drug absorption can occur from the vagina. For example, Patel *et al*[24] reported that plasma propranolol concentrations following vaginal dosing were significantly higher than those after peroral administration of an equivalent dose; a reflection, in part, of decreased first-pass biotransformation following vaginal absorption. Nonetheless, the notion persists that the vaginal epithelium is relatively impermeable to drugs. The widespread extemporaneous compounding of progesterone vaginal suppositories,[25,26] as well as the marketing of an intrauterine progesterone drug delivery system [Progestasert, *Alza*] have focused interest on systemic drug absorption following intravaginal administration. However, only limited reports of research on *in vitro* and *in vivo* aspects of vaginal absorption have appeared in the literature to date.

Bases

The USP lists the following as usual suppository bases: cocoa butter, glycerinated gelatin, hydrogenated vegetable oils, mixtures of polyethylene glycols of various molecular weights and fatty acid esters of polyethylene glycol.

Cocoa Butter—Theobroma oil, or cocoa butter, is a naturally occurring triglyceride. About 40% of the fatty acid content is unsaturated. As a natural material there is considerable batch-to-batch variability. A major characteristic of theobroma oil is its polymorphism, ie, its ability to exist in more than one crystal form. While cocoa butter melts quickly at body temperature, it is immiscible with body fluids; this may inhibit the diffusion of fat-soluble drugs to the affected sites.

If, in the preparation of suppositories, the theobroma oil is overheated, ie, heated to about 60°, molded and chilled, the suppositories formed will melt below 30°. The fusion treatment of theobroma oil requires maximum temperatures of 40 to 50° to avoid a change in crystal form and melting point. Theobroma oil, heated to about 60° and cooled rapidly will crystallize in an alpha configuration characterized by a melting point below 30°. The alpha form is metastable and will slowly revert to the beta form with the characteristic melting point approaching 35°. The transition from alpha to beta is slow, taking several days. The use of low heat and slow cooling allows direct crystallization of the more stable beta crystal form.

Certain drugs will depress the melting point of theobroma oil. This involves no polymorphic change although the net effect is similar. Chloral hydrate is the most important of those substances because its rectal hypnotic dose of 0.5 to 1.0 g will cause a substantial melting-point depression. This effect can be countered by addition of a higher melting wax, such as white wax or synthetic spermaceti. The amount to be added must be determined by temperature measurements. The effect of such additives on bioavailability also must be considered.

Water-Soluble or Dispersible Bases—Water-miscible suppository bases are of comparatively recent origin. The majority are comprised of polyethylene glycols or glycol-surfactant combinations. Water-miscible suppository bases have the substantial advantage of lack of dependence on a melting point approximating body temperature. Problems of handling, storage and shipping are simplified considerably.

Polymers of ethylene glycol are available as polyethylene glycol polymers (Carbowax, polyglycols) of assorted molecular weights. Suppositories of varying melting points and solubility characteristics can be prepared by blending polyethylene glycols of 1000, 4000 or 6000 molecular weight.

Polyethylene glycol suppositories are prepared rather easily by molding. The drug-glycol mixture is prepared by melting and then is cooled to just above the melting point before pouring into dry unlubricated molds. Cooling to near the melting point prevents fissuring caused by crystallization and contraction. Polyethylene glycol suppositories cannot be prepared satisfactorily by hand-rolling.

Water-miscible or water-dispersible suppositories also can be prepared using selected nonionic surfactant materials. Polyoxyl 40 stearate is a white, water-soluble solid melting slightly above body temperature. A polyoxyethylene derivative of sorbitan monostearate is water-insoluble but dispersible. In using surfactant materials the possibility of drug–base interactions must be borne in mind. Interactions caused by macromolecular adsorption may have a significant effect on bioavailability.

Examples of water-miscible suppository bases, devised by Zopf *et al*, are

Base 1

Polyethylene glycol 1000	96%
Polyethylene glycol 4000	4%

Base 2

Polyethylene glycol 1000	75%
Polyethylene glycol 4000	25%

Base 1 is low-melting and may require refrigeration; Base 2 is more heat-stable. Each is prepared conveniently by molding techniques.

Water-dispersible bases may include polyoxyethylene sorbitan fatty acid esters. These are either soluble (Tween, Myrj) or water-dispersible (Arlacel), used alone or in combination with other wax or fatty materials. Surfactants in suppositories should be used only with recognition of reports that such materials may either increase or decrease drug absorption.

Glycerinated Gelatin—Glycerinated gelatin usually is used as a vehicle for vaginal suppositories. For rectal use a firmer suppository can be obtained by increasing the gelatin content. Glycerinated gelatin suppositories are prepared by dissolving or dispersing the drug substance in enough water to equal 10% of the final suppository weight. Glycerin (70%) is then added and Pharmagel A or B (20%), depending on the drug compatibility requirements. Pharmagel A is acid in reaction, Pharmagel B is alkaline. Glycerinated gelatin suppositories must be formed by molding. The mass cannot be processed by hand-rolling. These suppositories, if not for immediate use, should contain a preservative such as methylparaben and propylparaben.

Preparation

Suppositories are prepared by rolling (hand-shaping), molding (fusion) and cold compression.

Rolled (Hand-Shaped) Suppositories—Hand-shaping suppositories is the oldest and the simplest method of pre-

paring this dosage form. The manipulation requires considerable skill, yet avoids the complications of heat and mold preparation.

The general process can be described as follows:

General Process

Take the prescribed quantity of the medicinal substances and a sufficient quantity of grated theobroma oil. In a mortar reduce the medicating ingredients to a fine powder or, if composed of extracts, soften with diluted alcohol and rub until a smooth paste is formed. The correct amount of grated theobroma oil then is added, and a mass resembling a pill mass is made by thoroughly incorporating the ingredients with a pestle, sometimes with the aid of a small amount of wool fat. When the mass has become plastic under the vigorous kneading of the pestle, it quickly is loosened from the mortar with a spatula, pressed into a roughly shaped mass in the center of the mortar and then transferred with the spatula to a piece of filter paper which is kept between the mass and the hands during the kneading and rolling procedure. By quick, rotary movements of the hands, the mass is rolled to a ball which immediately is placed on a pill tile. A suppository cylinder is formed by rolling the mass on the tile with a flat board, partially aided by the palm of the other hand, if weather conditions permit. The suppository "pipe" frequently will show a tendency to crack in the center, developing a hollow core. This occurs when the mass has not been kneaded and softened sufficiently, with the result that the pressure of the roller board is not carried uniformly throughout the mass but is exerted primarily on the surface. The length of the cylinder usually corresponds to about four spaces on the pill tile for each suppository, thus making the piece, when cut, practically a finished suppository except for the shaping of the point. When the cylinder has been cut into the proper number of pieces with a spatula, the conical shape is given it by rolling one end on the tile with a spatula, or in some cases even by shaping it with the fingers to produce a rounded point.

Compression-Molded (Fused) Suppositories—This method of suppository preparation also avoids heat. The suppository mass, such as a mixture of grated theobroma oil and drug, is forced into a mold under pressure, using a wheel-operated press. The mass is forced into mold openings, pressure is released, the mold removed, opened and replaced. On a large scale cold-compression machines are hydraulically operated, water-jacketed for cooling and screw-fed. Pressure is applied via a piston to compress the mass into mold openings.

Fusion or Melt Molding—In this method the drug is dispersed or dissolved in the melted suppository base. The mixture then is poured into a suppository mold, allowed to cool and the finished suppositories removed by opening the mold. Using this procedure, one to hundreds of suppositories can be made at one time.

Suppository molds are available for the preparation of various types and sizes of suppositories. Molds are made of aluminum alloy, brass or plastic and are available with from six to several hundred cavities.

Suppositories usually are formulated on a weight basis so that the medication replaces a portion of the vehicle as a function of specific gravity. If the medicinal substance has a density approximately the same as theobroma oil, it will replace an equal weight of oil. If the medication is heavier, it will replace a proportionally smaller amount of theobroma oil.

For instance, tannic acid has a density of 1.6 as compared with cocoa butter (see Table VII[27]). If a suppository is to contain 0.1 g tannic acid, then 0.1 g ÷ 1.6 or 0.062 g cocoa butter should be replaced by 0.1 g of drug. If the blank weight of the suppository is 2.0 g, then 2.0 − 0.062 g or 1.938 g cocoa butter is required per suppository. The suppository will actually weigh 1.938 g + 0.1 g or 2.038 g. Table VII indicates the density factor, or the density as compared with cocoa butter, of many substances used in suppositories.

It always is possible to determine the density of a medicinal substance relative to cocoa butter, if the density factor is not available, by mixing the amount of drug for one or more suppositories with a small quantity of cocoa butter, pouring the mixture into a suppository mold and carefully filling the mold with additional melted cocoa butter. The cooled suppositories are weighed providing data from which a working formula can be calculated as well as the density factor itself.

When using suppository bases other than cocoa butter, such as a polyethylene glycol base, it is necessary to know either the density of the drug relative to the new base or both the densities of the drug and the new base relative to cocoa butter. The density factor for a base other than cocoa butter is simply the ratio of the blank weight of the base and cocoa butter.

For instance, if a suppository is to contain 0.1 g tannic acid in a polyethylene glycol base, then 0.1 g ÷ 1.6 × 1.25 or 0.078 g polyethylene glycol base should be replaced by 0.1 g drug (the polyethylene glycol base is assumed to have a density factor of 1.25). If the blank weight is 1.75 g for the polyethylene glycol base, then 1.75 g − 0.078 g or 1.672 g of base is required per suppository. The final weight will be 1.672 g base + 0.1 g drug or 1.772 g.

When the dosage and mold calibration are complete the drug–base mass should be prepared using minimum heat. A water bath or water jacketing usually is used. The melted mass should be stirred constantly but slowly to avoid air entrapment. The mass should be poured into the mold openings slowly. Prelubrication of the mold will depend on the vehicle. Mineral oil is a good lubricant for cocoa butter

Table VII—Density Factors for Cocoa Butter Suppositories[27,28]

Medication	Factor
Acid, boric	1.5
Acid, benzoic	1.5
Acid, gallic	2.0
Acid, salicylic	1.3
Acid, tannic	1.6
Alum	1.7
Aminophylline	1.1
Aminopyrine	1.3
Aspirin	1.3
Barbital	1.2
Belladonna extract	1.3
Bismuth carbonate	4.5
Bismuth salicylate	4.5
Bismuth subgallate	2.7
Bismuth subnitrate	6.0
Castor oil	1.0
Chloral hydrate	1.3
Cocaine hydrochloride	1.3
Digitalis leaf	1.6
Glycerin	1.6
Ichthammol	1.1
Iodoform	4.0
Menthol	0.7
Morphine hydrochloride	1.6
Opium	1.4
Paraffin	1.0
Peruvian Balsam[a]	1.1
Phenobarbital	1.2
Phenol[a]	0.9
Potassium bromide	2.2
Potassium iodide	4.5
Procaine	1.2
Quinine hydrochloride	1.2
Resorcinol	1.4
Sodium bromide	2.3
Spermaceti	1.0
Sulfathiazole	1.6
Tannic acid	1.6
White wax	1.0
Witch hazel fluidextract	1.1
Zinc oxide	4.0
Zinc sulfate	2.8

[a] Density adjusted taking into account white wax in mass.

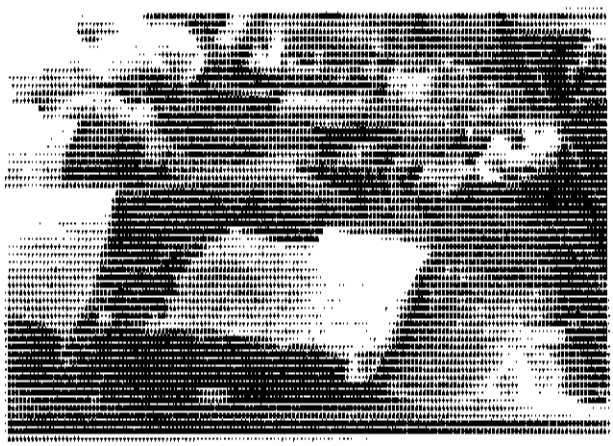


Fig 87-11. Removing cocoa butter suppositories from mold (courtesy, Webcon Div, Alcon).

suppositories. Molds should be dry for polyethylene glycol suppositories.

After pouring into tightly clamped molds the suppositories and mold are allowed to cool thoroughly using refrigeration on a small scale or refrigerated air on a larger scale. After thorough chilling any excess suppository mass should be removed from the mold by scraping, the mold opened and the suppositories removed. It is important to allow cooling time adequate for suppository contraction. This aids in removal and minimizes splitting of the finished suppository.

Packaging and Storage—Suppositories often are packaged in partitioned boxes which hold the suppositories upright. Glycerin and glycerinated gelatin suppositories often are packaged in tightly closed screw-capped glass containers. Though many commercial suppositories are wrapped individually in aluminum foil, or PVC-polyethylene strip-packaging is commonplace.

The most recent innovation in suppository manufacture is the procedure for molding the suppository directly into its primary packaging. In this operation the form into which the suppository mass flows consists of a series of individual molds formed from plastic or foil. After the suppository is poured and cooled the excess is trimmed off and the units are sealed and cut into 3s or 6s as desired. Cooling and final cartoning then can be carried out.

Suppositories with low-melting ingredients are best stored in a cool place. Theobroma oil suppositories, in particular, should be refrigerated.

Other Medicated Applications

Cataplasms (Poultices)

Poultices represent one of the most ancient classes of pharmaceutical preparations. A poultice, or cataplasm, is a soft, moist mass of meal, herbs, seed, etc, usually applied hot in cloth. The consistency is gruel-like, which is probably the origin of the word poultice.

Cataplasms were intended to localize infectious material in the body or to act as counterirritants. The materials tended to be absorptive, which, together with heat accounts for their popular use. None is now official in the USP. The last official product was Kaolin Poultice NF IX.

Pastes

Pastes are concentrates of absorptive powders dispersed (usually) in petrolatum or hydrophilic petrolatum. They are stiff to the point of dryness and reasonably absorptive in view of the petrolatum base. Pastes often are used in the treatment of oozing lesions where they act to absorb serous secretions. Pastes also are used to restrict the area of treatment by acting as an absorbent and physical dam.

Pastes adhere reasonably well to the skin, and are poorly occlusive. For this reason they are suited for application on and around moist lesions. The heavy consistency of pastes imparts a degree of protection and may, in some instances, make the use of bandages unnecessary. Pastes are less macerating than ointments.

Because of their physical properties pastes easily may be removed from the skin by the use of mineral oil or a vegetable oil. This is particularly true when the underlying or surrounding skin is traumatized easily.

An official paste is the conventional Zinc Oxide Paste; another is Triamcinolone Acetonide Dental Paste, for the specialized use the name implies.

Powders

Powders for external use usually are described as dusting powders. Such powders should have a particle size of not more than 150 μm, ie, less than 100-mesh, to avoid any sensation of grittiness which could irritate traumatized skin. Dusting powders usually contain starch, talc and zinc stearate. Absorbable Dusting Powder USP is comprised of starch treated with epichlorohydrin, with not more than 2.0% magnesium oxide added to maintain the modified starch in impalpable powder form; as it is intended for use as a lubricant for surgical gloves it should be sterilized (by autoclaving) and packaged in sealed paper packets.

The fineness of powders often is expressed in terms of mesh size, with impalpable powders generally in the range of 100- to 200-mesh (149–125 mm). Determination of size by mesh analysis becomes increasingly difficult as particle size decreases below 200-mesh.

Dressings

Dressings are external applications resembling ointments usually used as a covering or protection. Petrolatum Gauze is a sterile dressing prepared by adding sterile, molten, white petrolatum to precut sterile gauze in a ratio of 60 g of petrolatum to 20 g of gauze. Topical antibacterials are available in the form of dressings.

Creams

Creams are viscous liquid or semisolid emulsions of either the o/w or w/o type. Pharmaceutical creams are classified as water-removable bases and are described under *Ointments*. In addition to ointment bases, creams include a variety of cosmetic-type preparations. Creams of the o/w type include shaving creams, hand creams and foundation creams; w/o creams include cold creams and emollient creams.

Plasters

Plasters are substances intended for external application made of such materials and of such consistency as to adhere to the skin and attach to a dressing. Plasters are intended to afford protection and support and/or to furnish an occlusive and macerating action and to bring medication into close contact with the skin. Medicated plasters, long used for local or regional drug delivery, are the prototypical transdermal delivery system.

Plasters usually adhere to the skin by means of an adhesive material. The adhesive must bond to the plastic backing and to the skin (or dressing) with proper balance of cohesive strengths. Such a proper balance provides for re-

moval, ie, adhesive breakage at the surface of application thus leaving a clean (skin) surface when the plaster is removed.

Contraceptives

In the context of this chapter contraceptives are considered in the form of creams, jellies or aerosol foams intended for vaginal use to protect against pregnancy. Contraceptive creams and jellies are designed to melt or spread, following insertion, over the vaginal surfaces. These agents act to immobilize spermatozoa.

Creams and jellies for contraceptive use may contain spermicidal agents such as nonoxynol 9 or they may function by a specific pH effect. A pH of 3.5 or less has an appreciable spermicidal effect. It is important to note that a final *in situ* pH of 3.5 or less is required; thus, the dilution effect and pH change brought about by vaginal fluids must be considered. To achieve the proper pH effect and control, buffer systems composed of acid and acid salts such as lactates, acetates and citrates are used frequently. The user must, of course, be assured of the safety, lack of irritancy, acceptability and effectiveness of such products; also, detailed and specific information and instructions should be available to physicians.

References

1. Franz TJ: *J Invest Derm 64:* 194, 1975.
2. Scheuplein RJ, Ross LW: *Ibid 63:* 353, 1974.
3. Bartek MJ, La Bodde JA, Maibach HI: *Ibid 58:* 114, 1972.
4. Maibach HI, ed: *Animal Models in Dermatology*, Churchill Livingstone, Edinburgh, 110, 1975.
5. Wester RC, Maibach HI: *J Invest Derm 67:* 518, 1976.
6. Stoughton RB. In Maibach HI, ed: *Animal Models in Dermatology*, Churchill Livingstone, Edinburgh, 121, 1975.
7. Marahon MM, Callahan JF: *Ibid:* 36.
8. Lorenzetti OJ: *Ibid:* 212.
9. Maibach HI, ed: *Ibid:* 221.
10. Ostrenga J, Steinmetz C, Poulsen B: *J Pharm Sci 60:* 1175, 1971.
11. Walters KA. In Hadgraft J, Guy RH, eds: *Transdermal Drug Delivery*, Marcel Dekker, New York, 197, 1989.
12. Scheuplein RJ, Ross LW: *J Soc Cos Chem 21:* 853, 1970.
13. Christophers E, Kligman AM. In Montagna W, ed: *Advances in the Biology of Skin*, vol 6, Pergamon, Oxford, 163, 1965.
14. Sloan KB, Bodor N: *Int J Pharmaceut 12:* 299, 1982.
15. Banga AK, Chien YW: *J Control Rel 7:* 1, 1988.
16. Burnette RR. In Hadgraft J, Guy RH, eds: *Transdermal Drug Delivery*, Marcel Dekker, New York, 247, 1989.
17. Balsam MS, Sagarin E, eds: *Cosmetics Science and Technology*, 2nd ed, vol 1, Wiley-Interscience, New York, 205, 1972.
18. Lorenzetti OJ, Wernet TC: *Dermatologica 154:* 244, 1977.
19. Wood E, Wilson CG, Hardy JG: *Int J Pharm 25:* 191, 1985.
20. Tukker JJ, de Blaey CJ, Charbon GA: *Pharm Res 1:* 173, 1984.
21. Lowenthal W, Borzelleca JF: *J Pharm Sci 54:* 1790, 1965.
22. Parrott EL: *Ibid, 60:* 867, 1971.
23. Benziger DP, Edelson J: *Drug Metab Rev 14:* 137, 1983.
24. Patel LG, Warrington SJ, Pearson RM: *Br Med J 287:* 1247, 1983.
25. Roffo BD, Zimmer RA, Derewicz HJ: *Am J Hosp Pharm, 34:* 1344, 1977.
26. Allen LV, Stiles ML: *US Pharm 13(1):* 16, 1988.
27. Davis H: *Bentley's Text-Book of Pharmaceutics*, 7th ed, Bailliere, Tindall & Cox, London, 569, 1961.
28. Buchi J: *Pharm Acta Helv 20:* 403, 1940.

Bibliography

Cosmetic Ingredient Dictionary, 3rd ed, The Cosmetic, Toiletry and Fragrance Assoc, 1982.

Chien YW, ed: *Transdermal Controlled Systemic Medications*, Marcel Dekker, New York, 1987.

Chien YW: *Novel Drug Delivery Systems*, Marcel Dekker, New York, 1982.

Frost P, Gomez EC, Zaias N: *Recent Advances in Dermatopharmacology*, Spectrum Publ, New York, 1978.

Glaw B, delBlaey CJ, eds: *Rectal Therapy*, JR Prous, Barcelona, 1984.

Guy RH, Hadgraft J: In Maibach H, Lowe R, eds: *Models in Dermatology*, vol 2, 5 Karger, Basel, 170, 1985.

Hoover JE, ed: *Dispensing of Medication*, 8th ed, Mack Publ Co, Easton PA, 1976.

Maibach HI: *Animal Models in Dermatology*, Churchill Livingstone, Edinburgh, 1975.

Mier PD, Cotton DWK: *The Molecular Biology of Skin*, Blackwell, Oxford.

Montagna W, Parrakkal PF: *The Structure and Function of the Skin*, 3rd ed, Academic Press, New York, 1974.

Marples MJ: *The Ecology of the Human Skin*, Thomas, Springfield IL, 1965.

Walters KA: In Hadgraft J, Guy RH: *Transdermal Drug Delivery*, Marcel Dekker, New York, 197, 1989.