plot can be generated for an entire dosing regimen in a very short time. The next example illustrates such a calculation.

EXAMPLE. The following microcomputer program, written in BASIC, will solve Eq. (56) with an added loading dose for any time during a repetitive dosing regimen.

Program Listing 1

```
10  TEXT : HOME
20  VTAB 10: PRINT "MULTIPLE DOSE - PLASMA CONCENTRATION"
100  VTAB 15: PRINT "NEW PATIENT"
110  PRINT : PRINT "ENTER ELIMINATION HALF-LIFE IN HOURS": INPUT T
120 KEL = 0.69 / T
130  PRINT : PRINT "ENTER ABSORPTION HALF-LIFE IN HOURS": INPUT T
140 KA = 0.69 / T
150  PRINT : PRINT "ENTER VOLUME OF DISTRIBUTION AS FRACTION OF
     BODY WEIGHT": INPUT T
160  PRINT : PRINT "ENTER PATIENT WEIGHT IN KG": INPUT TT
170 VD = T * TT
200  PRINT : PRINT "NEW DOSE OR REGIMEN"
210  PRINT : PRINT "ENTER FRACTION OF DOSE ABSORBED": INPUT T
220  PRINT : PRINT "ENTER LOADING DOSE IN MG": INPUT TT
230 LD = T * TT
240  PRINT : PRINT "ENTER MAINTENANCE DOSE IN MG": INPUT TT
250 MD = T * TT
260  PRINT : PRINT "ENTER DOSING INTERVAL": INPUT T
270 TAU = T
280  PRINT : PRINT "ENTER TIME INTERVAL BETWEEN POINTS": INPUT T
290 STP = T
300  PRINT : PRINT "NEW TIME POINT"
310  PRINT : PRINT "ENTER NUMBER OF MAINTENANCE DOSES GIVEN":
     INPUT T
320 N = T
400  PRINT : PRINT "PLASMA CONCENTRATION AFTER"
410  PRINT N;" MAINTENANCE DOSES"
420  PRINT : PRINT "TIME (HR)","CONCENTRATION (MG/L)": PRINT
430  FOR TIME = 0 TO TAU STEP STP
440 E1 = 0: IF (1 -  EXP (KA * TAU)) = 0 THEN  GOTO 460
450 E1 = (LD *  EXP ( - N * KA * TAU) + MD * (1 -  EXP ( - N * KA
     * TAU)) / (1 -  EXP ( - KA * TAU))) *  EXP ( - KA * TIME)
460 E2 = 0: IF (1 -  EXP (KEL * TAU)) = 0 THEN  GOTO 480
470 E2 = (LD *  EXP ( - N * KEL * TAU) + MD * (1 -  EXP ( - N *
     KEL * TAU)) / (1 -  EXP ( - KEL * TAU))) *  EXP ( - KEL *
     TIME)
```

FRESENIUS EXHIBIT 1012

EXAMPLE. Calculate C_p at quarter-hourly intervals for the first 24 hr of the kanamycin dosing regimen described in the first example in this subsection. The entry data are as follows:

Half-life for elimination	= 2.3 hr
Half-life for absorption	= 0.2 hr
Fraction of body weight equal to volume of distribution	= 0.27
Subject weight	= 74 kg
Fraction of the dose absorbed	= 1.0
Loading dose (D_0)	= 800 mg
Maintenance dose (d_0)	= 800 mg
Dosing interval (T)	= 6 hr
Time interval	= 0.25 hr
Number of maintenance doses (n)	= 0–4

The results of the calculation are shown in Fig. 15. Note that C_{max} and C_{min} are identical with the values calculated in the aforementioned example.

C. Dosage Regimen Adjustment in Renal Failure

The influence of impaired renal or liver function on the rate of elimination of a drug was mentioned previously, and a method for predicting the half-life of kanamycin in patients with varying degrees of renal impairment was described earlier (see Sec. VI). If patients with im-

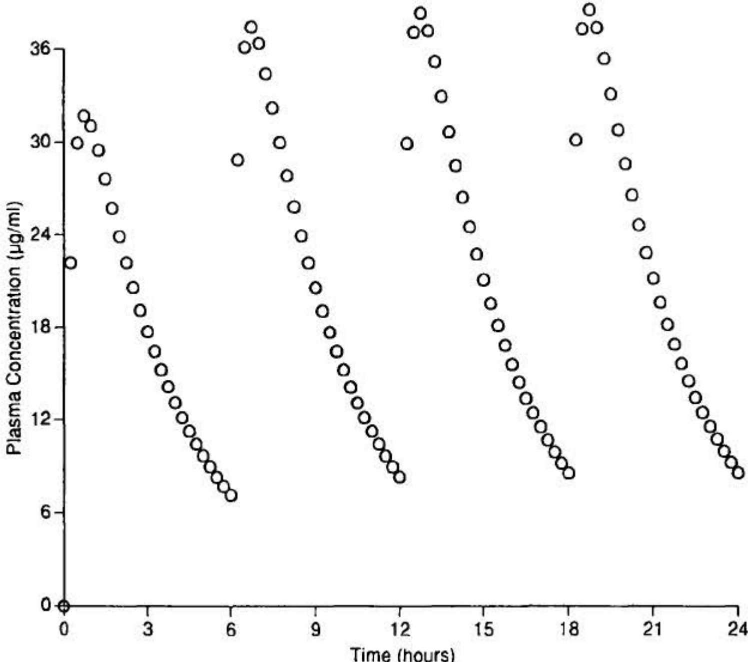


Fig. 15 Plot of kanamycin plasma concentrations (circles) versus time following multiple IM injections, calculated with Eq. (56).

paired renal function are given a normal-dosing regimen of kanamycin, they will soon build up toxic plasma concentrations of the drug. However, they can be dosed safely and effectively by adjusting the dosing regimen in accordance with the predicted elimination half-life. *Dosing regimen adjustment involves changing the dose or the dosing interval.*

Alteration of the Dosing Interval

The kanamycin package insert recommends that a dose of 7.5 mg/kg be administered every three half-lives, and that for patients with impaired renal function, the half-life in hours can be estimated by multiplying the serum creatinine level (mg/100 ml) by 3 (see also Sec. VI). Thus the dosing interval for a renal patient should be nine times serum creatinine level (mg/100 ml).

EXAMPLE. A patient (74 kg) has a serum creatinine level of 6 mg/100 ml (elimination half-life ≈ 18 hr); therefore, the dosing regimen would be 7.5 mg/kg IM every 54 hr. Figure 16 shows the results of this dosing regimen in terms of the plasma concentrations of kanamycin produced in the renal patient (solid line) and a normal patient (dashed line). In both cases, the C_{max} values are below the MTC of 35 μg/ml, but the C_{min} values fall below the MEC of 10 μg/ml (Sec. IX.A).

For the normal patient, the time during which the plasma concentrations are below 10 μg/ml is only about 4 hr for each dosing interval; but for the renal patient, this time extends to about 18 hr. This gives the renal patient inadequate therapy for long time periods and can foster the development of strains of bacteria that are resistant to kanamycin.

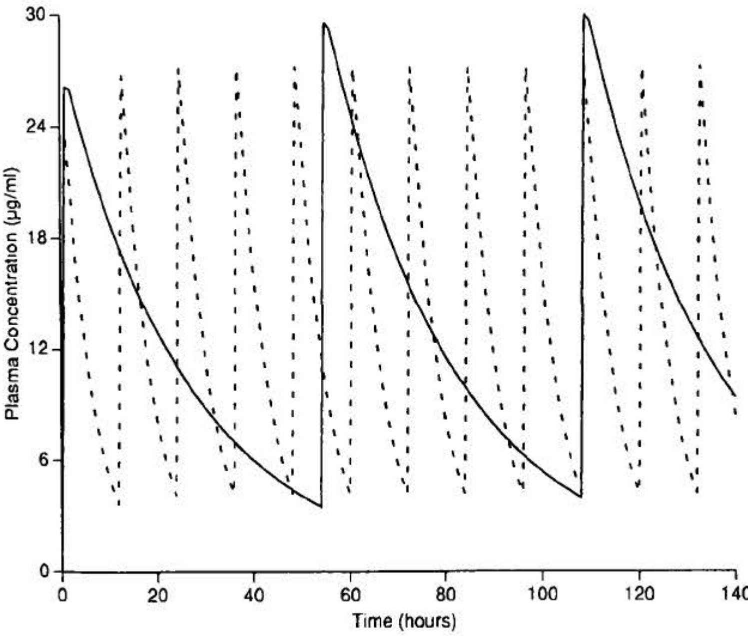


Fig. 16 Plot of kanamycin plasma concentration versus time following multiple IM injections in a normal patient (dashed line) and a renal patient (solid line), calculated with Eq. (56) and $T = 9 \times$ serum creatinine level (mg/100 ml) (*altered-dosing interval* approach).

FRESENIUS EXHIBIT 1012
Page 102 of 293

Alteration of the Dose

In a previous example the "maintenance dose every half-life" regimen was proposed for kanamycin because it maintained plasma concentrations between the MTC and MEC. The same regimen could be employed for IM injections of kanamycin in renal patients, but the logistical problems that arise in the clinic when the maintenance doses must be given at odd time intervals (e.g., 18 hr) make this regimen somewhat impractical. It would be much better if the maintenance doses could be administered at the same time other medication is given (i.e., every 4, 8, or 12 hr).

The same loading dose (in milligrams per kilogram) can be given to all patients regardless of their renal function because the loading dose is determined by the volume of distribution (or body weight) and not by the rate of elimination. However, if a normal-dosing interval is used, the amount of drug eliminated over one interval (T) will be much less for a patient with renal failure than for a patient with normal renal function. As a result, the maintenance doses must be reduced to replace only that amount of drug lost during the preceding dosing interval.

EXAMPLE. The dosing regimen recommended for normal adults in the kanamycin package insert is 7.5 mg/kg every 12 hr. The insert also states that the half-life of kanamycin in a normal adult is about 4 hr. What should the 12-hr maintenance dose be for the individual in the previous example?

For a person whose half-life is 18 hr (k_{el} = 0.0385 hr^{-1}), the amount remaining at the end of a 12-hr–dosing interval is

$$\log \frac{A_0}{A} = \frac{k_{el}t}{2.30}$$

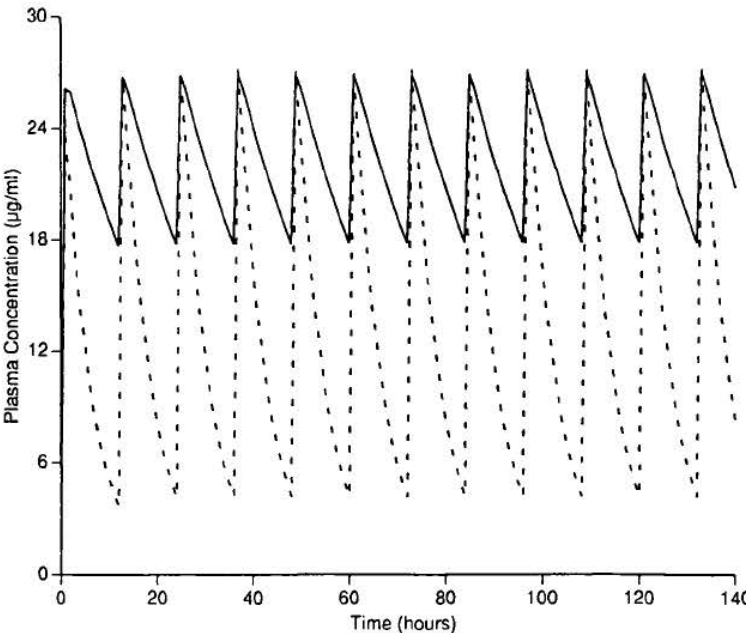


Fig. 17 Plot of kanamycin plasma concentration versus time following multiple IM injections in a normal patient (dashed line) and a patient with renal impairment (solid line) calculated with Eq. (56) and maintenance dose = amount lost over each 12-hr–dosing interval (*altered dose* approach).

FRESENIUS EXHIBIT 1012
Page 103 of 293

$$\log \frac{7.5}{A} = \frac{0.0385 \times 12}{2.30}$$

$$A = 4.72 \text{ mg/kg}$$

The amount lost over the 12-hr interval is

$$7.5 \text{ mg/kg} - 4.72 \text{ mg/kg} = 2.78 \text{ mg/kg}$$

Therefore, the maintenance dose = 2.78 mg/kg every 12 hr. The loading dose (same as normal) = 7.5 mg/kg.

The results of this dosing regimen in terms of plasma concentrations are shown as a solid line in Fig. 17. The dashed line in Fig. 17 shows the plasma concentrations that would be produced in a normal patient on a normal-dosing regimen (7.5 mg/kg every 12 hr). Figure 17 shows that the administration of 2.8 mg/kg maintenance doses every 12 hr to the aforementioned renal patient is a convenient dosing regimen that produces C_{min} plasma levels above the MEC and C_{max} plasma levels below the MTC.

REFERENCES

1. R. E. Notari, *Biopharmaceutics and Clinical Pharmacokinetics: An Introduction*, 4th Ed., Marcel Dekker, New York, 1987.
2. M. Gibaldi and D. Perrier, *Pharmacokinetics*, 2nd Ed., Marcel Dekker, New York, 1982.
3. J. G. Wagner, *Fundamentals of Clinical Pharmacokinetics*, Drug Intelligence Publications, Hamilton, IL, 1975.
4. J. G. Wagner, *Biopharmaceutics and Relevant Pharmacokinetics*, Drug Intelligence Publications, Hamilton, IL, 1971.
5. G. A. Portmann, Pharmacokinetics, in *Current Concepts in the Pharmaceutical Sciences: Biopharmaceuticals* (J. Swarbrick, ed.), Lea & Febiger, Philadelphia, 1970.
6. W. J. Westlake, The design and analysis of comparative blood-level trials, in *Current Concepts in the Pharmaceutical Sciences: Dosage Form Design and Bioavailability* (J. Swarbrick, ed.), Lea & Febiger, Philadelphia, 1973.

5

The Effect of Route of Administration and Distribution on Drug Action

Svein Øie and Leslie Z. Benet
University of California, San Francisco, California

I. THE DOSE-EFFICACY SCHEME

When a health practitioner administers (or "inputs") a dose of drug to a patient, usually the ultimate goal is solely directed to the usefulness of the drug under abnormal conditions. That is, the drug must be efficacious and must be delivered to its site of action in an individual experiencing a particular physiological anomaly or pathological state. Pharmaceutical scientists, on the other hand, concentrate their attention to solving problems inherent in drug delivery to deliver the optimal dose to the site(s) of action.

The general pathway a drug takes from residence in a dosage form until its clinical utility is depicted in Fig. 1. Ideally, the drug should be placed directly at the site of action, as illustrated by the stippled arrow in Fig. 1, to maximize the effect and minimize side effects relating to unwanted responses at sites other than the target tissue. However, delivery directly to the site of action is more often than not, impractical or not possible. Instead, we have to settle for the most convenient routes of delivery. This is illustrated by the solid arrows in Fig. 1. That is, the drug is placed directly in the vascular systems or in close proximity to some biological membrane through which the drug can traverse to reach body fluids or the vascular system. The delivery system is generally designed to release the drug in a manner that is conducive to this passage through the membrane. Previous chapters have discussed how drug delivery systems may be optimized in terms of dissolution in the fluids surrounding the membrane to allow the desired rate of passage through the membrane. Subsequent chapters will deal with specific drug delivery systems and their optimization. Once the drug has passed through the membrane and into the blood stream adjacent to the site of absorption, a general distribution of the drug will take place throughout the biological system. As pointed out in Chapter 3, the degree of dilution (referred to as the apparent volume of distribution) will dictate the initial concentration of drug in the general circulation, as sampled from a peripheral vein. Usually, the dose of the drug administered to the patient was chosen to give sufficiently high

FRESENIUS EXHIBIT 1012
Page 105 of 293

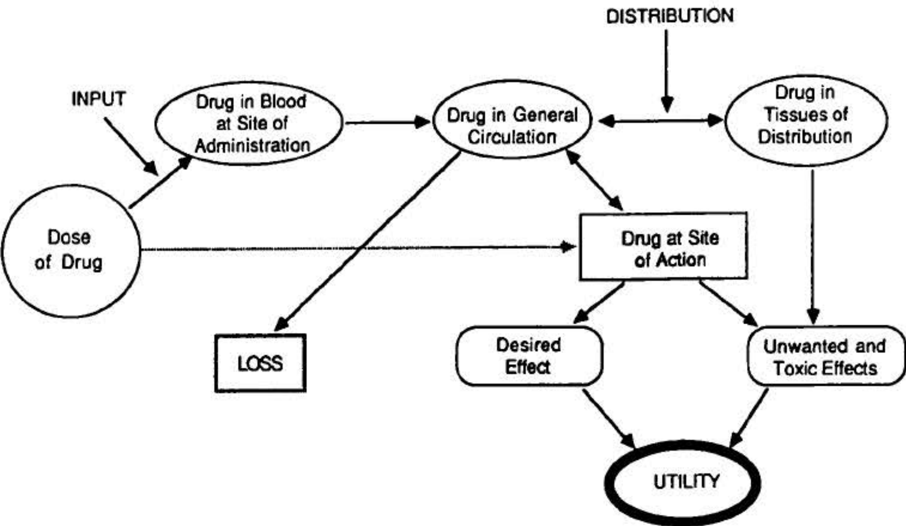


Fig. 1 A schematic representation of the dose–efficacy relationship for a drug.

blood levels so that an adequate quantity of the drug would reach the site of action. The rate of input needed to achieve adequate levels of the drug at the site of action is influenced not only by the distribution and general elimination in the body, but may also be modified by the loss processes that are unique to a specific route of administration. This chapter will deal primarily with the distribution and loss processes that result uniquely from the physiological parameters inherent in the use of a particular route of administration.

Unfortunately, no drug is yet so specific that it interacts with only the target site in the target tissue, and will not give rise to hyperclinical activity. Too much drug at the wrong place or too high a concentration at the right place may result in unwanted or toxic effects. Thus the practitioner must determine the usefulness of any dose of a drug from a particular drug delivery system by balancing the efficacy achieved from the clinical effect against the toxic reactions observed.

Most drug delivery systems achieve the required drug levels at the site of action as a result of attaining adequate blood levels in the general circulation (see Fig. 1, solid arrows). This process is followed because of the ease with which present drug delivery systems can "input" drugs into the general circulation and the inherent difficulties in delivering the drug selectively to a relatively inaccessible site (e.g., pituitary gland). In addition, for many compounds, the exact site of action is still unknown. However, when the site of drug action is sufficiently defined, Fig. 1 illustrates the advantage of delivering the drug directly to the site of action. By direct administration to the active site, a lower dose could be used to achieve the clinical effect because the drug no longer is diluted or eliminated en route. As a result, drug concentrations at unwanted sites of action could be kept to a minimum; in addition, clinically effective levels at the site of action might be attained much more rapidly, since the process of distribution throughout the entire body could be avoided. One should not forget that, in addition to the obvious clinical advantage of direct administration, there is also an economical one. By delivering the drug to the site of action, the amount of drug needed is much smaller than by more traditional delivery methods. This is particularly important for many of the newer recombinant

compounds that can be very expensive. Much work is currently being carried out in an attempt to achieve such a selectiveness as that described in Fig. 1.

II. PHYSIOLOGICAL CONSIDERATIONS FOR THE VARIOUS ROUTES AND PATHWAYS OF DRUG INPUT

A. Drug Input at or Close to the Site of Action

Figure 2 illustrates a number of sites where drug delivery systems have historically been used to input drug directly to its site of action [1,2]. Various classic dosage forms were developed to take advantage of these input sites: eye, ear, and nose drops; inhalation, oral, topical, and vaginal aerosols; topical solutions, creams, and ointments; and rectal solutions, enemas, and suppositories. Each of the sites for local drug administration requires specific formulation to allow the drug to remain at the site of application for a sufficient length of time to allow the drug to penetrate through the particular membrane(s) so that it can reach the actual site of action adjacent to the site of application. For example, some ophthalmic preparations may be given to elicit a superficial anti-infective effect, such as treatment of an inflammation of the conjuctiva. Thus, only topical effects are desired, and there is no need for the drug to penetrate into the eyeball. Formulation of such products would be quite different than formulation of a drug delivery system for which the drug must be absorbed into the interior of the eye to produce

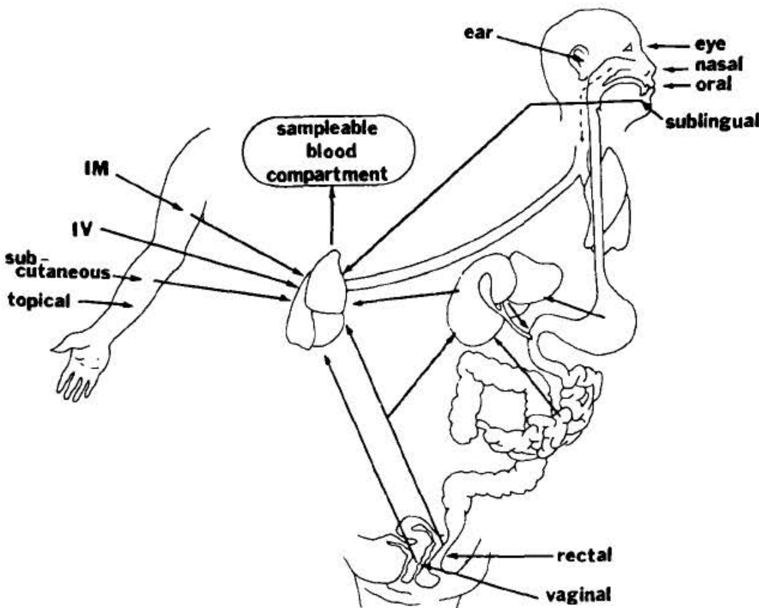


Fig. 2 Various routes and pathways by which a drug may be "input" into the body. The position of one lung is distorted to emphasize that the lungs are in an excellent position for cleansing the blood. The diagram is especially useful in explaining the first-pass effect following oral dosing, for which drug absorbed from the small intestine or stomach must first pass through the liver and, therefore, is subject to metabolism or biliary excretion before reaching the sampleable blood. (From Ref. 2.)

a response, such as miotics, mydriatics, anti-inflammatory drugs that act in the anterior segment of the eye, and, occasionally, drugs for treatment of infections. A detailed description of the factors involved in the development of such ophthalmic preparations, as well as the particular physiological characteristics of the eye, will be presented in Chapter 15. Similar types of design problems arise for many of the other sites that are traditionally treated by direct local application. One of the most difficult problems facing the formulator is that the behavior of the diseased tissue may be different from that for healthy individuals, and it may also change over the course of treatment. For example, diseased skin is often more permeable than healthy skin; therefore, the drug may disappear faster from the site of administration than desired, and the effect will be less than expected. Should the formulation be designed to accommodate this phenomenon, one must be mindful of the fact that as the pathological condition improves, the absorption may also change.

Although the classic dosage forms mentioned earlier can be used to put drug directly into the site of action, many of them have a degree of "messiness" that prevents good patient acceptance and adherence. Not only is there an initial psychological barrier that must be overcome, but the general public has an aversion to taking drugs by routes other than oral. There is, in addition, a general dislike for sticky creams, drippy drops, greasy ointments, and the like. Over the last two decades much work has been directed toward developing more acceptable delivery systems than the traditional ones. Emphasis has been placed on long-acting drug delivery systems that may be more convenient, since they would only require self-administration once a week or possibly at even longer intervals.

A large number of new devices have been developed, and new ones are constantly being investigated. Plastic disks for placement in the eye (similar to a contact lens) that slowly release drug into the humoral fluid; drug-impregnated plastic rings or loops that when placed in the uterus will release controlled amounts of contraceptive agents; bioadhesive tablets or disks that can be placed buccally, nasally, or vaginally for local release; hydrogels for slow release in the eye are examples of such new delivery systems that input drug directly to the site of action.

In a more ambitious move, many groups have also embarked on site-specific delivery to less accessible sites than those given in Fig. 2. Although numerous experimental systems have been designed, few have reached the clinical stage. The simplest and most direct method when a specific target organ can be located is cannulation (direct access port). A catheter is placed in an appropriate artery or vein. If a vein is used, the catheter has to reach the organ, or otherwise the drug will be flowing away from the target tissue, be diluted with blood from the rest of the body, and be not different than a systematic intravenous administration. Catheters can also be placed in the peritoneum, the bladder, and in the cerebrospinal fluid. A drug can now be administered directly into the desired tissues at a rate that can be well controlled. Although catheter delivery is a direct method, it is limited in that it is essentially restricted to inpatient use. Use of implants in the desired tissue, or a drug carrier (e.g., liposomes, nanoparticles, and such) that will either home in on the desired tissue by specific receptors, or release their content at the desired site by an external stimulus (e.g., magnetic fields, light, current), are drug systems currently being explored for target-specific delivery.

Although the method of direct delivery is a very attractive one, it also has its regulatory problems. Benet [1] has noted that assessing the bioavailability of this system can create difficulties because the manufacturer may not be able to devise a control procedure that can measure drug concentration at the site of action. For example, the extent and rate of availability of an orally administered drug can easily be assessed by measuring blood levels, whereas for a drug input into a site of action, significant blood levels would indicate distribution away from that site. Frequently, significant blood levels of a drug that is administered at a site of action (such as a topical preparation, an eye drop, a nasal insufflation, or an antibiotic that acts

FRESENIUS EXHIBIT 1012
Page 108 of 293

on intestinal flora) indicate either a poor drug delivery system or substantial overdosing. For this class of drug delivery systems, clinical efficacy necessarily has to serve as the best measurement of drug availability and dosage form efficacy.

B. Drug Input into the Systemic Circulation

The overwhelming majority of existing drugs are, however, given by general routes; that is, by routes that do not deliver the drug directly to the site of action. These modes of drug input rely on a passive delivery of drug through distribution by the vascular system. The most commonly accepted method is oral administration. As will be discussed later, oral administration is not ideal, as one needs to be concerned about whether the drug can be destroyed in the stomach, in the gastrointestinal fluid, in its passage through the gut wall, through the liver, or simply not be absorbed in time before it is expelled from the gastrointestinal tract. Several alternative routes of delivery are being used or are being developed to diminish these potential losses. The advantages and problems inherent in the individual routes of administration will now be discussed.

Parenteral Administration: Intravascular

Of the routes of input depicted in Fig. 2, intravenous (IV) administration yields one of the fastest and most complete drug availabilities. However, intra-arterial injections might be employed when an even faster and more complete input of drug to a particular organ is desired. By administering the drug through an artery, the total drug delivered will enter the organ or tissue to which the artery flows. Intravenously administered drug will first be diluted in the venous system as the venous blood is pooled in the superior and inferior vena cava. It then enters the heart, and is subsequently pumped to the lung before it can enter the arterial system and reach the target organ(s). In addition, the fraction of the drug reaching a desired site is dependent on the fraction of the arterial blood flow reaching that site. Additional drug can reach the target tissue only by being recirculated from the other organs. In comparison with intra-arterial administration, IV administration reaches the target slower, and initially at a lower concentration. Although intra-arterial injections appear superior, they are infrequently used because they are considered much more dangerous than IV administration. Intra-arterial administration has been associated with patient discomfort, bleeding, and thrombosis.

In addition to the dilution factor resulting from mixing with larger volumes of blood after intravenous administration, one also needs to consider the possibility of temporary or permanent loss of drug during its passage through the lung. The position of the lungs in Fig. 2 has been distorted to emphasize the point that the lungs are in an excellent strategic position for cleansing the blood, since all of the blood passes through the lungs several times a minute. Apart from their respiratory function and the removal of carbon dioxide from the pulmonary circulation, the lungs serve other important cleansing mechanisms, such as filtering emboli and circulating leukocytes, as well as excretion of volatile substances. The lungs also have metabolic capacity [3] and may serve as a metabolic site for certain drugs [4] or as an excretory route for compounds with a high vapor pressure. The lungs can also act as a good temporary storage site for a number of drugs, especially basic compounds, by partitioning of the drug into lipid tissues, as well as serving a filtering function for particulate matter that may be given by IV injection. Accumulation of lipophilic compounds and filtering of any compounds in solid form can be viewed as a temporary clearing or dilution of the drug, as it will eventually leach back into the vascular system. Thus the lung serves as a dampening or clearing device, that is not present following intra-arterial injection. Drugs given by the IV route may, therefore, not necessarily be completely available to the site of action, since a certain fraction of the drug could be

FRESENIUS EXHIBIT 1012
Page 109 of 293

eliminated by the lung before entering into the general circulation [5]. This might be called a "lung first-pass effect."

The foregoing concepts may be visualized by referring to Fig. 3. In this figure one can readily see the difference between intra-arterial and intravenous administration of drugs. Let us assume that compartment n is the target tissue. Administration into any vein (i.e., into any of the efferent arrows on the left-hand side of the figure) would lead the drug to the heart and, from there, to the lung. Drug that enters the lung can leave by only one of two routes, as illustrated in Fig. 4: by the blood that leaves the lung, or by being eliminated. The result is that there is a competition between the two routes for the drug, and the greater the ability of the lung to eliminate the drug in comparison with the pulmonary blood flow, the more drug will be extracted. If we assume that the pulmonary blood flow is Q_P and that the intrinsic elimination clearance of the organ is CL_{Int}, and no plasma protein binding occurs ($Cu = C_{out}$), then the extraction ratio can be expressed as

$$E = \frac{C_{out}\, CL_{Int}}{(C_{out}\, CL_{Int}) + (C_{out}\, Q_P)} = \frac{CL_{Int}}{CL_{Int} + Q_P} \tag{1}$$

In perfusion models, as depicted in Fig. 3, it is assumed that distribution into and out of the organ is perfusion rate-limited such that drug in the organ is in equilibrium with drug concentration in the emergent blood [6]. The intrinsic clearance of an organ is different from the value we normally think of as the clearance of the organ. The *clearance of the organ* is defined as the rate of loss in relation to the incoming concentration, whereas the *intrinsic clearance* is defined as the rate of loss in relation to the organ concentration (or exiting concentration). In addition, it is also clear that, of the drug that escapes elimination in the lung, only a small fraction goes to compartment n, the rest is distributed to other organs. Drugs that enter these organs will be exposed to elimination in these organs and must necessarily recirculate through the heart and lungs before they again have the opportunity to reach compartment n.

Parenteral Administration: Depot

The other parenteral routes depicted in Fig. 2, intramuscular (IM) and subcutaneous (SC) injections, may also be considered in terms of Fig. 3. Drug absorbed from the IM and SC sites into the venous blood will return to the heart and pass through the lungs before being distributed to the rest of the body. However, there will be an initial lag between the time when the drug is injected and when it enters the circulation. Thus, the kinetics for drugs administered by these parenteral routes would be expected to show a decreased *rate* of availability and may also show a decreased *extent* of availability in comparison with intravenous administration, if loss processes take place at the site of injection. For example, we could consider that drug is now injected directly into compartment m in Fig. 3 and that this compartment is the muscle. The rate at which the drug leaves the muscle will depend primarily on blood flow in relation to the size (apparent volume of distribution) of the organ.

Evans and co-workers [7] measured resting human muscle blood flow through the gluteus maximus, vastus lateralis, and deltoid muscles. Deltoid muscle blood flow was significantly greater than gluteus muscle blood flow, with vastus being intermediate. Because the two sites most commonly used for IM injections are the deltoid and the gluteus muscles, we might expect to see differences in rates of drug absorption following injection to these sites. Lidocaine is one drug that has been investigated for its effect in response to the site of injection [8,9]. Deltoid injection gave higher peak levels than lateral thigh injection which, in turn, gave higher levels than gluteal injection. Schwartz et al. [9] demonstrated that therapeutic plasma levels for a particular lidocaine dose were reached only when the deltoid injection site was used. Evans et al. [7] concluded, "This demonstrates that the site of injection can influence the

FRESENIUS EXHIBIT 1012
Page 110 of 293

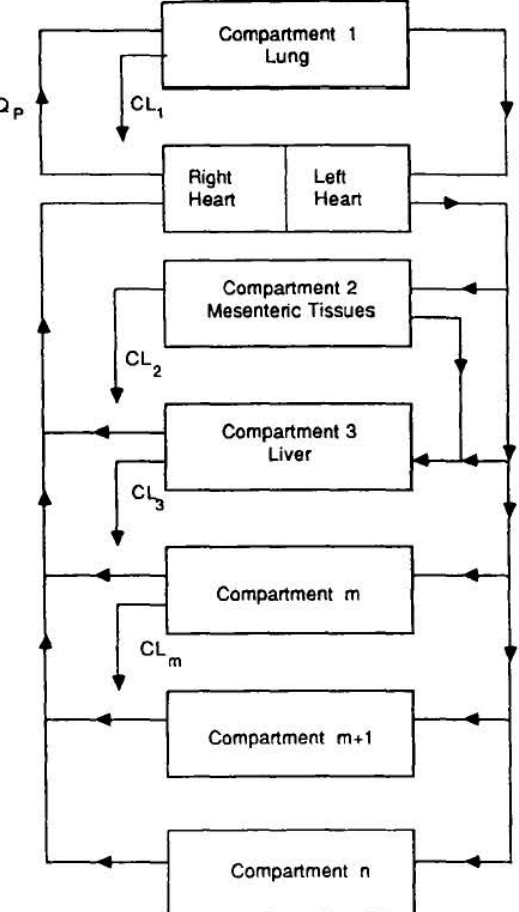


Fig. 3 The body depicted as a physiological perfusion model. Compartment m must be considered as a summation of the individual tissues that metabolize the drug and compartment m + 1 through n as noneliminating tissues.

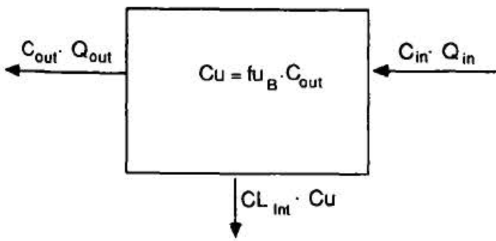


Fig. 4 Flow model of an eliminating organ. The drug enters the body by the organ blood flow, CL_{in} and is immediately mixed in the organ. The drug leaves the organ by either being eliminated ($CL_{int}Cu$) or by the exiting blood flow.

FRESENIUS EXHIBIT 1012
Page 111 of 293

plasma level achieved and that the deltoid muscle should be used to achieve therapeutic blood levels as rapidly as possible." Likewise, if a sustained or a prolonged release is desired, this would more readily be achieved by injection into a lower blood flow muscle, such as the gluteus.

Loss processes may also account for a decrease in the extent of availability following an IM injection. This can be visualized by assuming that the dose is injected into compartment m, which, as depicted in Fig. 3, is capable of eliminating the drug. As shown in Fig. 4, the drug can leave the tissue only by one of two routes, either by the blood leaving the organ, or by being eliminated by metabolism in the muscle. In addition, the drug that leaves the site of administration will also be subject to the additional distribution and elimination in the lung, similar to intravenous administration. In other words, drug given by intramuscular administration may be not only further delayed in its distribution to the target organ, but may also show a decreased extent of distribution to the organ, in comparison with the intravenous dose. For example, degradation can take place in the muscle, as shown by Doluisio et al. [10] for ampicillin. These workers found that only 77–78% of an IM dose of ampicillin sodium solution was absorbed, as compared with the IV solution. The most likely explanation is that the drug may have been decomposed chemically or enzymatically at the injection site. In addition, temporary losses may also occur. For example, intramuscular doses of phenytoin result in a marked decreased rate and extent of absorption in comparison with IV or oral doses. Wilensky and Lowden [11] demonstrated that this could be due to precipitation of the drug as crystals in the muscle. Although these crystals eventually dissolve, the drug is essentially lost during a normal dosing interval.

Oral Administration

First-Pass Effect. Metabolism in the Gastrointestinal Fluids and Membranes. When a dosage form is administered by the oral route, drug particles come in contact with varying pH solutions, different enzymes, mucus, gut flora, and bile, all of which may contribute to decreasing the extent of availability by degradation, binding, or sequestering mechanisms. These factors, as well as the possiblity of drug metabolism in the intestinal membrane itself, have been well covered in Chapter 2 and will not be discussed here.

Hepatic Metabolism: Linear Systems. As depicted in both Figs. 2 and 3, drug that is absorbed from the gastrointestinal tract must pass through the liver before reaching the sampleable circulation and the rest of the body. Thus, if a drug is metabolized in the liver or excreted into the bile, some of the active drug absorbed from the gastrointestinal tract will be inactivated by hepatic processes before the drug can reach the general circulation and be distributed to its site(s) of action. An exception would be if the liver itself were the target organ, as we then would have to contend with only losses in the gastrointestinal tract and in the gut wall before reaching the site of action.

For many drugs, the fraction of the dose eliminated on the first pass through the liver is substantial. The fraction eliminated is often referred to as the hepatic extraction ratio, designated herein as E_H. Many drugs are known or suspected to have a high hepatic extraction ratio. A short list of some of the better-known compounds is given in Table 1. The hepatic first-pass phenomenon is not restricted to any particular pharmacological or chemical group of drug substances and the foregoing list includes acids, bases, and neutral compounds.

The available fraction (F) of an oral dose appearing in the sampleable blood circulation will, therefore, be governed by not only the extent of drug absorbed from the gastrointestinal tract, as discussed in Chapter 4, but also by the fraction metabolized in the gut membranes (E_G) and the fraction metabolized or excreted into the bile following passage through the liver (E_H), where E_G and E_H are the extraction ratios for the gut and liver, respectively. If it is

Table 1 Drugs Suspected to Have High Hepatic Extraction Ratios

Acetylsalicylic acid	Alprenolol	Aldosterone
Cocaine	Desipramine	Doxorubicin
Fluorouracil	Isoproteranol	Imipramine
Lidocaine	Lorcainide	Morphine
Nitroglycerin	Prazepam	Propranolol

assumed that a drug is completely absorbed from the gastrointestinal tract and not degraded during passage through the gut membranes, this would be equivalent to an injection of the drug into the hepatic portal vein. Under these conditions the unmetabolized fraction (F_H) of an oral dose appearing in the sampleable blood circulation would equal $1 - E_H$. If the drug elimination in the liver follows first-order kinetics, Rowland [12] has shown that F_H, the fraction of the oral dose that is available following the liver first pass, may be related to liver blood flow (Q_H) and the hepatic clearance for an IV dose of the drug ($CL_{H,iv}$).

$$F_H = 1 - E_H = 1 - \frac{CL_{H,iv}}{Q_H} \tag{2}$$

This hepatic first-pass effect, as discussed by Boyes et al. [13], is well illustrated in Fig. 5. Identical doses of lidocaine hydrochloride were injected into dogs (beagles) by an exponential infusion process. The upper curve describes the mean levels found when the drug was infused into a peripheral vein, whereas the lower curve describes the lidocaine levels following hepatic portal vein infusion (thereby eliminating any effects caused by gastrointestinal degradation either in the fluids of the tract or in the intestinal membrane). The area under the curve (AUC)

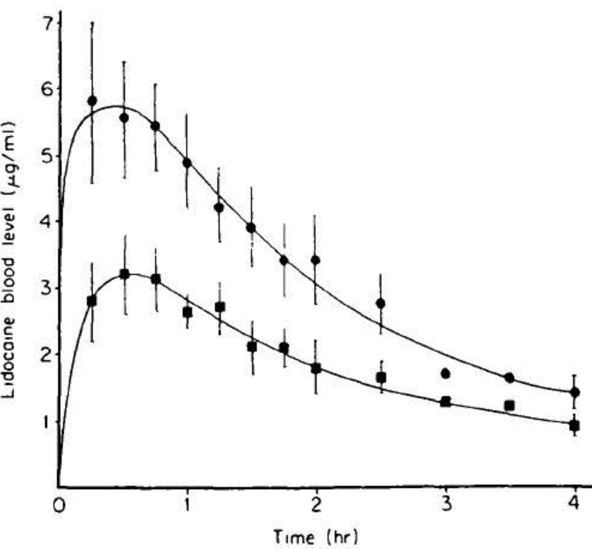


Fig. 5 Mean blood levels of lidocaine hydrochloride in five beagles after exponential IV infusion through a peripheral vein (●) and in the portal vein (■). Vertical bars represent standard errors of the mean. (From Ref. 13.)

measurements for the two curves show that the extent of bioavailability following portal vein infusion is only 60% of that found following infusion into a peripheral vein (i.e., $F_H = 0.6$). This "oral" availability could be predicted for a drug such as lidocaine for which drug elimination occurs predominantly in the liver when the hepatic clearance—calculated by dividing the dose by the area under the peripheral vein infusion curve—and the hepatic blood flow are substituted into Eq. (2). An even greater first-pass effect is found for oral administration of lidocaine to humans, $F_H = 0.25–0.48$ [14].

Thus, if the hepatic clearance for a drug is largely relative to the hepatic blood flow, the extent of availability for this drug will be low when it is given by a route that yields first-pass effects. The decrease in availability is a function of only the anatomical site from which absorption takes place, and no amount of dosage form redesign can improve the availability. Of course, therapeutic blood levels can be reached by this route of administration if larger doses are given, but the health practitioner and pharmacist must be aware that levels of the drug metabolite may increase significantly over that seen following IV administration.

Lidocaine is not normally administered by the oral route, but many of the drugs listed earlier are routinely given orally. For these drugs, analysis of Eq. (2) leads to the conclusion that small variations in plasma or blood clearance of a drug throughout a population may yield significant differences in availability when the drug is given by a route subject to significant first-pass effects. For example, data from a study of Shand et al. [15] for oral and IV administration of propranolol in five men are shown in Table 2. Although the clearance following an IV dose of the drug (column 6) varies by only 67% from the smallest to the largest, the oral availability (column 5) varies by 275%. As would be expected from Eq. (2), the oral availability decreases as the IV clearance increases.

Hepatic Metabolism: Nonlinear Systems. The discussion of the first-pass effect has thus far assumed linear first-order kinetics. Under this condition, the hepatic extraction will be independent of the rate of drug availability. That is, no matter when a drug molecule is absorbed from the gastrointestinal tract and at whatever dose administered, the hepatic extraction and the extent of availability [see Eq. (2)] for that drug will remain constant. This would not be true for a drug for which saturation of the hepatic enzymes is a possibility. Under such a condition, the extraction ratio would vary depending on the concentration of drug in the hepatic portal vein. If the concentration were high, the hepatic enzymes would become partially saturated, and a large amount of drug would pass through the liver without being metabolized (i.e., E_H would be low). Saturable metabolism during the first passage through the liver is not

Table 2 Peak Plasma Levels and Areas Under Plasma Concentration Time Curves Following Oral and Intravenous Administration to Men

	Propranolol, 80 mg fasting orally		Propranolol, 10 mg IV		
Subject	Peak (ng/ml)	Area (ng/ml·hr)	Area (ng/ml·hr)	$\frac{AUC_{Oral}}{AUC_{IV}} \cdot \frac{10}{80} \cdot 100$	Clearance, IV (ml/min)
OF	212	1400	292	60	570
DS	100	480	220	30	756
GY	94	510	200	32	833
JC	45	290	183	20	909
JF	36	220	175	16	950

Source: After Ref. 15.

an unusual occurrence even for a drug that, after intravenous administration, fails to show saturable metabolism. The main reason is that drug absorbed from the gastrointestinal tract is diluted to only a small degree before it enters the liver (in the portal blood only), and the concentrations can be relatively high. On the other hand, intravenously administered drug is generally quickly distributed to individual tissues, and the concentration reaching the liver during the elimination process is sufficiently low that no saturation of the metabolism is achieved. However, if the concentration in the hepatic portal vein were low, either because of administration of a low dose or because of very slow absorption of the drug, then the enzymes would not be saturated, and most of the drug could be metabolized on passage through the liver (i.e., E_H would approach 1). In a similar manner, metabolism in the gastrointestinal membranes could also be saturated or not, depending on the dose and the rate of absorption (i.e., E_G would vary from 0 to 1).

Salicylamide appears to be a drug for which the first-pass extraction may be dose-dependent. When a 300-mg oral dose of salicylamide is administered as a solution, the area under the plasma concentration time curve is less than 1% of the area seen following a similar IV dose [16]. Even following a 1-g oral dose, most of the drug is found in the systemic circulation as the inactive glucuronide and sulfate conjugates. However, when a 2-g dose was given to a particular subject (Fig. 6), the area under the plasma concentration time curve for unmetabolized salicylamide increased dramatically (> 200 times increase over that seen for a 1-g dose). This very large increase is probably related to saturation of enzymes in the gastrointestinal mucosa as well as in the liver. Saturation of the hepatic enzymes not only increases the extent of oral availability, but a sufficient amount of unchanged drug is allowed to enter the systemic circulation, thereby decreasing the rate of drug metabolism during the postabsorptive phase as well. This phenomenon is due to enzyme saturation. Thus, the extraction ratio [see Eq. (1)] and, thereby, the drug clearance, change continuously during both the absorption and elimination processes. Under these conditions AUC measurements can no longer be used to deter-

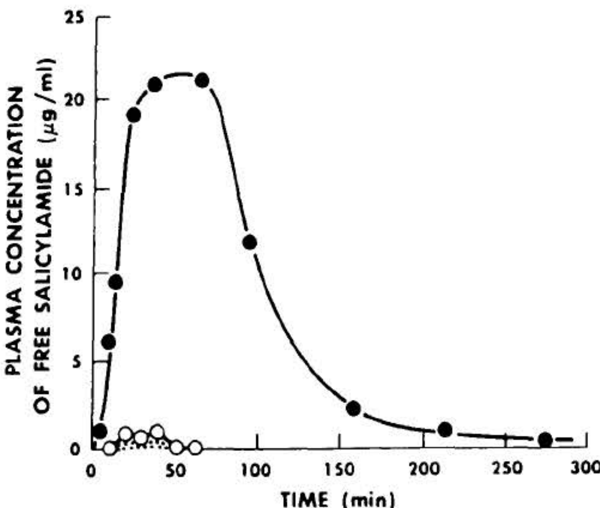


Fig. 6 Comparison of plasma concentrations of intact salicylamide when given as a 1-g (○) and a 2-g (●) dose in solution to the same subject. Dotted lines show the plasma concentrations of a 0.5-g dose and 0.3-g dose. (From Ref. 16.)

mine the extent of drug availability (as was previously described in Chapter 3 for linear systems.)

The effects of stomach emptying on drug availability are discussed in Chapters 2 and 4, relative to processes within the gastrointestinal system. In addition, Benet [1] has interpreted data for *p*-aminobenzoic acid (PABA) as reflecting an example for which the rate of drug absorption modified by changes in stomach emptying causes changes in the extent of drug availability owing to saturation of first-pass metabolism. Table 3 represents a situation in which the drug PABA is completely absorbed from the gastrointestinal tract, yet there is a decrease in the extent of drug available to the systemic circulation as a function of food [17].

The data for the oral solution, single-dose studies suggest that all of the drug is being absorbed, since the total amount of PABA found in the urine (both as unchanged drug and acetyl metabolite) equals the dose administered. It appears that the decreasing fraction of the acetyl metabolite found in the urine with increasing doses would indicate that a saturable metabolic process was operable. For the second series of studies listed in Table 3, the oral solution dose remains constant at 1 g, while increasing amounts of sweet cream are added to the solution. With increasing amounts of fat, there is a concomitant increase in the percentage of dose excreted as the acetyl metabolite. This may be explained by assuming that the fat decreases the rate of stomach emptying, causing the drug to be emptied from the stomach more slowy and to be absorbed over a longer time period. If plasma concentration were maintained at lower levels by slowing absorption, the metabolic site would not reach the same degree of saturation and a greater fraction of the metabolite should appear in the urine. Even if gastric absorption of PABA did occur, owing to retention of the drug in the stomach, the absorption rate should be significantly lower than that seen for the oral solutions without fat, thereby maintaining plasma concentrations of PABA at lower levels. The prolonged administration of the smaller oral and IV doses in the third part of Table 3 yields high levels of metabolite in the urine, which is consistent with the saturable enzyme hypothesis (i.e., when plasma concentrations are maintained at a low level, extensive metabolism will occur).

Table 3 Extent of Urinary Excretion of *p*-Aminobenzoic Acid (PABA) and Its Acetyl Metabolite as a Function of Route of Administration and Ingestion of Fat

Route	Dose Na-PABA to 61-kg man (g)	Total PABA in urine as percentage of dose in 24 hr	Acetyl-PABA in urine as percentage of total PABA excreted in 24 hr
Oral solution			
Single dose	1	103	51
	2	103	47
	4	102	36
	8	102	30
Fat added to oral dose			
60 g sweet cream	1	95	76
90 g sweet cream	1	104	83
120 g sweet cream	1	99	90
Prolonged administration			
10 oral doses given every 30 min	0.365	95	97
Intravenous infusion, 270 min	0.365	90	95
Intravenous bolus	1	102	51

Source: Ref. 17.

Biliary Excretion. The effects of significant hepatic extraction as a result of biliary secretion, with or without metabolism, would be expected to follow the same principles just outlined for hepatic metabolism. In fact, a whole class of compounds that serve as biliary contrast agents for radiological examination depend on significant first-pass biliary secretion to be effective.

Several studies in rats have shown that certain acidic and basic compounds can be actively secreted into the bile. Thus, one might expect to see saturation of the biliary excretion process, although data in humans describing this phenomenon have not, as yet, been reported for orally dosed drugs.

Oral Dosing Without a First-Pass Effect. If a drug is not metabolized in the gut wall or the liver and if the drug is not subject to biliary excretion, there will be no first-pass effect following oral dosing. In addition, with drugs for which the hepatic clearance is significantly less than hepatic blood flow [see Eq. (2)], the hepatic extraction will be negligible (i.e., $F_H \approx 1$). Most orally administered drugs used today fall into this latter category. Examination of Fig. 2 indicates that another form of oral administration, sublingual dosing, may avoid first-pass metabolism as well as the degradation process that may occur in the gastrointestinal fluids. This route of administration has been used predominantly in dosing organic nitrates in patients experiencing angina.

Rectal Administration

As can be seen in Fig. 2, the first-pass effect can be partially avoided by rectal administration. The capillaries in the lower and middle sections of the rectum drain into the inferior vena cava, thus bypassing the liver. However, suppositories tend to move upward in the rectum into a region where veins (such as the superior hemorrhoidal vein) drain predominantly into the portal circulation [18]. In addition, there are extensive anastomoses between the middle and superior hemorrhoidal veins. Thus, Schwarz [19] has suggested that only about 50% of a rectal dose can be assumed to bypass the liver and its first-pass extraction. Again referring to Fig. 2, we can see that absorption of drugs from the vagina would bypass the first-pass effect, although almost no drugs where systemic levels are desired have been formulated using this route of administration.

Other Routes of Administration

It is clear from the foregoing discussion that none of the methods traditionally used for systemic administration are ideal. Intravenous administration requires hospital staff and, therefore, is rarely used in an outpatient setting. Intramuscular and subcutaneous doses may be metabolized at the site of administration, may be significantly delayed, and may run the risk of infections and meet with public reluctance for self-administration, although subcutaneous administration has been met with great success in insulin-dependent diabetics. Oral administration must run through a series of elimination sites before the drug can enter the general circulation. Rectal administration can also be exposed to hepatic first-pass and must also contend with a public reluctance in its usage. Because of these factors, several other routes for systemic administration of drugs have been explored over the years. Significant acceleration in these area has occurred as a response to the increased interest in the use of peptide and protein drugs made by recombinant techniques. Peptides and proteins are notoriously prone to degradation in the gastrointestinal tract, and the only modes of administration have been parenteral administrations. This limits the use by outpatients, and great efforts are being made to develop alternative rontes of delivery.

Nasal Administration. The nasal mucosa is relative permeable to small molecular weight compounds. The most notorious example is cocaine. Cocaine that is snorted is both rapidly

and extensively absorbed. Small peptides have also been successfully administered nasally, although the bioavailability is low. However, where the availability is not critical, nasal administration of peptides has been successful. The best example is calcitonin, which shows an activity after nasal administration, similar to that seen after intravenous administration, although the plasma concentrations achieved after nasal delivery were lower [20]. This is a compound that has a relatively large therapeutic index for which a variable bioavailability is not critical.

During the absorption process through the nasal lining, the drug has to cross not only lipophilic barriers, but it must also pass through the nasal epithelia, which have a significant capability to metabolize drugs. Nasally administered drug can, in addition to this loss by metabolism, also be removed by mucous flow and ciliary movement and be swallowed. Nasal administration can, under circumstances during which significant amounts are swallowed, be thought of as being similar to administering the drug in part as an oral dose. To increase the absorption by the nasal epithelia, the use of bioadhesive drug delivery forms and many exciting absorption enhancer techniques are being explored. Chemicals that disrupt the lipophilic membranes as well as opening the tight junction between the cells have resulted in a dramatic increase in the availability. The opening of the tight junctions is particularly interesting, as the compounds pass between the cells, thereby avoiding exposure to the metabolizing enzymes in the nasal epithelia. Whether these absorption enhancement techniques are the ways of the future as a means to reduce first-pass elimination, or if, in the end, they will be judged to invoke too much tissue damage for continuous use, still needs to be evaluated.

Transdermal Absorption. If transdermal delivery of drugs to the systemic circulation is to be successful, it must mimic subcutaneous injection in terms of yielding minimal first-pass skin elimination. The main barrier to absorption is the thick lipophilic keratin layer of the skin. Few nonlipophilic compounds penetrate the skin to a sufficient degree that this mode of delivery can be used for systematic absorption without modification. Highly lipophilic drugs, on the other hand, penetrate the skin with relative ease, although the absorption usually takes a long time. Several drugs are successfully given by dermal administration (e.g., nitroglycerin, scopolamine, nicotine, progesterone). To increase the absorption of hydrophilic compounds, including peptides and proteins, strategies similar to those described previously for nasal absorption have been studied (i.e., use of hydration and chemical enhancers). In addition, use of nonchemical enhancement methods (i.e., iontophoresis) is also being explored. This technique uses a low electric charge to force fluid and solutes to cross the skin. Although it was introduced more than 200 years ago [21], it is only recently that the method appears to have reached a practical stage [22]. These methods not only allow the drug to bypass the gastrointestinal tract and the liver, but they also may be employed to achieve very attractive long-term sustained delivery.

Pulmonary Inhalation. Although aerosol preparations now serve primarily as a convenient drug delivery system that can input drug directly to its site of action, new interest has recently been generated in this delivery system as a potential route for systemic administration of drug. The lung has a relatively large surface area and is relatively permeable to lipophilic compounds and, to some degree, even to protein [23]. Several barriers to absorption by the lungs do exist. The barrier to absorption is greatest in the upper bronchi and decreases in the alveolie. In the upper bronchi the mucus is relatively thick, the surface area small, and ciliar movement tends to move impacted particles up the bronchi and into the esophagus. The particle size of therapeutic inhalation aerosols determines the site of deposition in the lungs and, thus, the clinical effectiveness of a particular formulation. Particles that are too large will impact in the upper bronchi, and those too small will not readily impact on the wall of the lung and, therefore, will simply be exhaled. Particle sizes in the order of 0.3–1 μm are usually considered to be

most effective. Sciarra [24] has suggested that almost all drugs given by IV injection can be reformulated into a suitable aerosol, provided that the drug is capable of being deposited in the respiratory tract and is nonirritating. However, the total availability of the pulmonary route will, to a large degree, depend on how much drug is deposited in the lung and, again, how much metabolized during first pass. It is expected that the first-pass bioavailability will be lower than that seen after intravenous administration because the drug has to pass the epithelial cell layer before it can reach the general circulation. From intravenous administration, only the amount of drug that is actually taken up by the lung tissue will be exposed to metabolism and exhalation.

III. DRUG DISTRIBUTION

Figure 1 indicates that distribution will take place as the drug reaches the general circulation. This will dilute the drug and influence the levels at the site of action. Thus, an understanding of drug distribution is critical in designing appropriate drug dosage regimens. This has led to the determination of "apparent" volumes of distribution (as discussed in Chapter 3), which can be used to relate the amount of drug in the body (or in a hypothetical compartment) to a measured plasma or blood concentration. The volume of distribution is a function of four major factors: (a) the size of the organs into which the drug distributes; (b) the partition coefficient of drug between the organ and the circulating blood; (c) the blood flow to the distributing organs; and (d) the extent of protein binding of the drug both in the plasma and in various tissues.

A. Organ Size, Blood Flow, and Partition Coefficient

A particular organ in the body may act as a site of distribution, or as both a site of distribution and elimination. The relative importance of the various organs as storage or elimination sites depends on how fast the drug gets to each organ and how much space or volume is available to hold the drug. Table 4 presents a compilation of the volumes and blood flows of the different regions of the human body for a standard man, as compiled by Dedrick and Bischoff [25] using the mean estimates of Mapleson [26].

The various regions of the body are listed in decreasing order relative to blood flow per unit volume of tissue (adrenals highest and bone cortex lowest). This value essentially describes how fast a drug can be delivered to a body region per unit volume of tissue, and reflects the relative rates in which tissues may be expected to come to equilibrium with the blood. How much drug can be stored or distributed into a tissue will depend on the size of the tissue (volume) and the ability of the drug to concentrate in the tissue (i.e., the partition coefficient between the organ and blood, $K_{O/B}$). For example, the blood flow per unit volume of thyroid gland (see region C in Table 4) is one of the highest in the body, whereas the gland itself is quite small. Thus, if partition of the drug between the thyroid and blood were approximately 1, we would expect to see that the drug in the tissue would rapidly come into equilibrium with that in the blood, but that relatively little drug would be found in the thyroid. However, for certain drugs containing iodine moieties, $K_{O/B}$ is enormous, and a significant amount of drug will distribute into this small gland relatively rapidly. In addition, Table 4 lists the volume of blood contained within each tissue and believed to be in equilibrium with the tissue. Thus, the volume of the thyroid in Table 4 is considered to be 20 ml of tissue and 49 ml of blood. Note that total volume of all tissues in column 3 is 70 liters, including the 5.4 liters of blood volume. This blood volume is broken down into 1.4 liters of arterial blood (which is listed in the last column as being in equilibrium with the air in the lungs) and 4 liters of venous blood (which

Table 4 Volumes and Blood Supplies of Different Body Regions for a Standard Man[a]

Tissue	Reference letters	Volume (liters)	Blood flow (ml/min)	Blood flow (ml/100 ml)	Volume of blood in equilibrium with tissue (ml)
Adrenals	A	0.02	100	500	62
Kidneys	B	0.3	1240	410	765
Thyroid	C	0.02	80	400	49
Gray matter	D	0.75	600	80	371
Heart	E	0.3	240	80	148
Other small glands and organs	F	0.16	80	50	50
Liver plus portal system	G	3.9	1580	41	979
White matter	H	0.75	160	21	100
Red marrow	I	1.4	120	9	74
Muscle	J	30	300/600/1500	1/2/5	185/370/925
Skin				1/2/5	18/37/92
Nutritive	K	3	30/60/150		
Shunt	L		1620/1290/300	54/43/10	
Nonfat subcutaneous	M	4.8	70	1.5	43
Fatty marrow	N	2.2	60	2.7	37
Fat	O	10.0	200	2.0	123
Bone cortex	P	6.4	≈0	≈0	≈0
Arterial blood	Q	1.4			
Venous blood	R	4.0			
Lung parenchymal tissue	S	0.6			
Air in lungs	T	2.5 +half			1400[b]
		Tidal volume			999/795/185[c]
Total		70.0[d]	6480		5400

[a]Standard man = 70-kg body weight, 1.83-m² surface area, 30–39 years old.
[b]Arterial blood.
[c]Skin-shunt venous blood.
[d]Excluding the air in the lung.

is in equilibrium with tissues A through O). Since different muscle masses throughout the body receive different blood flows (as discussed in Sec. II.B, with reference to drug input following IM injection), Mapleson [26] lists only a range for this tissue as well as for the skin. Dedrick and Bischoff [25] suggest an average blood flow value of 3.25 ml/(100 ml tissue × min) for tissues J and K, corresponding to average total flows of 980 and 98 ml/min for muscle and skin, respectively. Note that total blood flow in column 4 corresponds to the cardiac output, 6.48 liters/min.

When discussing drug distribution, it is often convenient to lump various tissue regions into general categories. For example, following an IV bolus injection, the heart, brain, liver, and kidneys achieve the highest and earliest drug concentrations, with equilibrium between these tissues and blood being rapidly achieved. Thus, Dedrick and Bischoff [25] have combined these tissues and other well-perfused regions (see A through H in Table 4) into a well-perfused compartment that they designate as viscera. Similarly, regions I, J, K, and M are lumped into a less well-perfused compartment, the lean tissues, whereas poorly perfused regions N and O are designated as the adipose compartment. Blood flows, volumes, and such, for these lumped compartments can be calculated by summing the appropriate terms in Table 4. By using a

FRESENIUS EXHIBIT 1012
Page 120 of 293

perfusion model containing these three lumped compartments and a blood compartment, Bischoff and Dedrick [27] were able to describe thiopental concentrations in various tissues as shown in Fig. 7.

Levels in the dog liver, representative of the visceral tissues, are already at a maximum by the time the first sample is taken, since very rapid equilibrium is achieved between these tissues and the blood. Drug uptake into the less well-perfused skeletal muscle, representative of the lean tissue, is slower—peaking at about 20 min, but still achieving apparent distribution equilibrium between 1 and 2 hr. Uptake into the poorly perfused adipose tissue is even slower. In fact, peak levels in this tissue have not even been reached by the time the last samples are taken.

Since the site of action for the barbiturates is the brain, we might expect the pharmacological action to correspond to the time course of thiopental concentrations in the viscera which, in turn, would be reflected by blood levels, since a rapid equilibrium is attained between viscera and blood. Although the pharmacological action may terminate quickly (within an hour) owing to decreased blood levels, traces of the drug may be found in the urine for prolonged periods (days) owing to accumulation in the fatty tissues. Note that the partition coefficient for the drug between tissues and blood is greater than 1 for all three tissue groups (i.e., at distribution

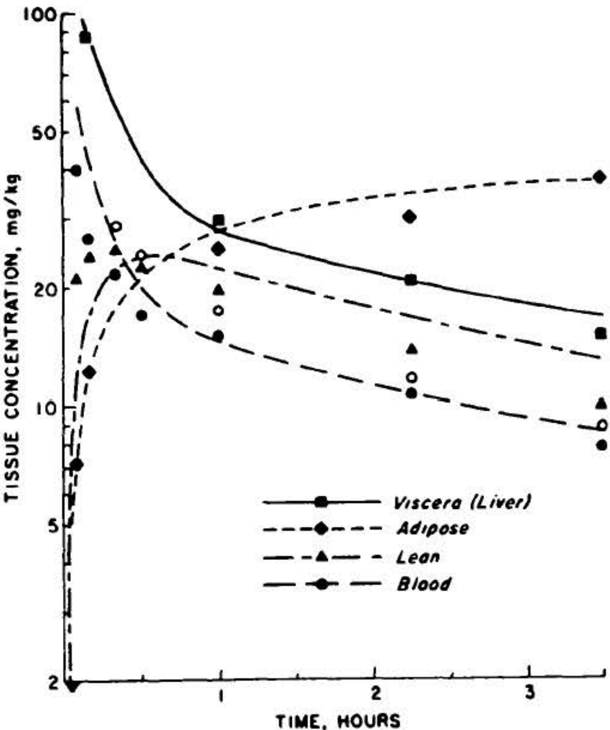


Fig. 7 Thiopental concentrations in various tissues following 25-mg/kg IV bolus doses. Solid symbols indicate data in dogs; the open circles are from data in humans. Lines correspond to predicted values in various tissues using a perfusion model containing compartments corresponding to the blood, viscera, lean, and adipose tissues. (From Ref. 26.)

equilibrium tissue concentrations greater than blood concentrations) but that the $K_{O/B}$ for adipose tissue is very large. This can be deducted from the fact that drug continues to distribute into adipose tissue, even when the concentration in the fat is significantly greater than the blood concentration. At $3^1/_2$ hr, most of the lipid-soluble thiopental left in the body is in the fat, and at later times, this percentage may even increase before distribution equilibrium is reached. At these later times the removal rate of the barbiturate from the body will be controlled by the slow movement of drug out of the fatty tissue, as a result of the high partition into the fat and the low blood flow to this region.

B. Protein Binding

One of the major determinants of drug distribution is the extent of protein binding. Many drugs bind to plasma proteins, mainly to albumin and α_1-acid glycoprotein (orosomucoid), but sometimes to lipoproteins, various globulins, and specific binding proteins as well. Up to the present time, most studies of drug protein binding have examined the interaction between drugs in the plasma and plasma albumin and plasma α_1-acid glycoprotein. Total plasma albumin for a 70-kg man is about 120 g. Total interstitial albumin is approximately 156 g [28]. Thus almost 60% of total albumin in the body is found outside the plasma. The total extracellular α_1-acid glycoprotein is approximately 3 g, of which approximately 45% is located extravascularly [29]. α_1-Acid glycoprotein has also been found intracellularly and exists as a membrane-bound form [30], but its total cellular amount is unknown. The extracellular amount of α_1-acid glycoprotein can increase significantly in many diseases (inflammation, infections, cancers) as well during trauma. The increase is variable and usually averages a twofold change, although increases up to fivefold have been reported. Binding to tissue components is also assumed to be important for the overall distribution, although direct evidence is difficult to obtain. However, albumin and α_1-acid glycoprotein represent only approximately 1% of the "dry tissue" in the body, of which a large fraction is proteins. One might, therefore, expect that the binding of drugs to tissue would affect drug distribution significantly and frequently much more than the binding of drugs to plasma proteins. The binding of drugs to plasma proteins has been studied extensively, primarily because the experiments can be easily carried out. Tissue-binding studies do not have this advantage, and thus knowledge of the qualitative and quantitative aspects of the binding of drugs to tissue components is poorly understood. The partition coefficients between body organs and blood, $K_{O/B}$, discussed in the previous section will be considerably influenced by binding. It is probably for this reason that attempts to correlate drug distribution, drug action, and membrane transport with oil/water partition coefficients have succeeded only infrequently, necessitating the use of different organic solvents in different correlations to obtain at least an approximate rank order correlation.

Effect on Distribution

The extent of distribution to an organ is usually expressed as the apparent volume of distribution of the organ, V_t. Because the blood concentration is used as a reference, the apparent volume of an organ, V_t, can usually be expressed in the following way:

$$V_t = V_T\, K_{O/B} = V_T \left(\frac{fu}{fu_T}\right) \quad (3)$$

where V_T is the physical volume of the tissue, fu is the unbound fraction of drug in the blood and fu_T is the unbound fraction of drug in the tissue. This relationship indicates that the partition coefficient $K_{O/B}$ essentially reflects the strength of binding to the tissues in comparison with the binding to blood proteins, which will be particularly true for hydrophilic compounds.

Decreased binding to proteins in a tissue will increase the value of fu_T, as the fraction of drug in the tissue that is not bound is now increased, and the apparent volume of the tissue decreases. Similarly, if the binding to blood proteins is decreased, the relative affinity for the tissue increases, and the apparent volume of distribution is increased. Often, volumes of distribution and protein binding are defined in terms of plasma measurements, rather than in blood, as used here. Such an approach is correct as long as one defines explicitly the reference fluid of measurement. We prefer to use blood measurements throughout, because, as described in the next section, the relationship of clearance to organ elimination capacity must be defined in terms of the total flow to each organ.

Effect on Clearance

The *clearance* of an organ can be viewed as the volume of blood completely cleared of drug per unit time. Given this definition, the blood clearance of an organ will be equal to the blood flow through the organ multiplied by the extraction ratio (equal to the fraction of the blood completely cleared of drug). From Eq. (1) we, therefore, obtain

$$CL = Q\,E = Q\left(\frac{CL_{Int}}{Q + CL_{Int}}\right) = \frac{Q\,CL_{Int}}{Q + CL_{Int}} \tag{4}$$

Equation (1) was derived assuming no protein binding, whereas, in fact, CL_{Int} should be defined in terms of the maximum ability of the organ to remove ***unbound*** drug because only unbound drug can bind to enzymes and excretory molecules. By including the unbound fraction of drug in Eq. (4) one obtains

$$CL_{organ} = \frac{Q\,fu\,CL_{Int}}{Q + fu \cdot CL_{Int}} \tag{5}$$

where fu is the fraction of drug unbound in the blood. [Note that all the discussions in this chapter have related clearances and flows to blood concentrations (i.e., drug concentrations in the actual transporting fluid). Blood concentrations must be used when any physiological interpretation is placed on the clearance process. However, it is easier to carry out studies measuring plasma concentrations and plasma protein binding. These measured parameters may be converted to blood values if one knows the blood/plasma concentration ratio.]

When the intrinsic clearance of an organ is very high compared with blood flow, such that $fu\,CL_{Int} >> Q$, the extraction ratio approaches 1, and $CL_{organ} \rightarrow Q$. Thus, according to Eq. (5), if a drug is cleared exclusively in the liver with an extraction ratio approaching unity, organ clearance will be very sensitive to changes in liver blood flow and essentially independent of blood binding [31]. Likewise, for drugs eliminated in the kidney by active-transport processes, elimination appears to be independent of the extent of blood binding. Under these conditions, kidney blood flow becomes the rate-limiting step.

However, for many drugs, blood flow is significantly greater than the intrinsic organ clearance ($Q >> fu\,CL_{Int}$) and then Eq. (5) reduces to

$$CL_{organ} = fu\,CL_{Int} \tag{6}$$

indicating that organ clearance is dependent on binding in blood and is proportional to the unbound fraction of drug in blood. This phenomenon is valid for both hepatic and renal elimination. For example, the glomerulus in the kidneys will filter approximately 120 ml of plasma per minute. High molecular weight compounds (i.e., most of the plasma proteins) will not be filtered. Likewise, drugs bound to these proteins will be retained. In other words, one could view the filtration process as being filtration of plasma water and of the small molecular weight compounds dissolved therein. Plasma water contains only unbound drug. Increases in binding

will lower the fraction of the drug in plasma in unbound form and, therefore, decrease the filtration clearance. For hepatic elimination, data for several compounds have been published over the years that show a correlation between the clearance and the unbound fraction in plasma [32]. We expect similar relationships between the organ clearance and the unbound fraction to be also valid for other organs.

Effect on Availability

Should changes in protein binding have a significant effect on first-pass availability? Combining Eqs. (2) and (5) we obtain:

$$F = 1 - E = 1 - \frac{\text{CL}}{Q} = 1 - \left[\frac{fu\ \text{CL}_{\text{Int}}}{Q + fu \cdot \text{CL}_{\text{Int}}}\right] = \frac{Q}{Q + fu \cdot \text{CL}_{\text{Int}}} \quad (7)$$

This indicates that the first-pass availability is a function of organ flow, protein binding, and intrinsic clearance of the organ. When $fu\ \text{CL}_{\text{Int}} >> Q$ (i.e., when we have relatively large extraction ratios), the first-pass bioavailability is equal to

$$F = \frac{Q}{fu\ \text{CL}_{\text{Int}}} \quad (8)$$

Under this circumstance, the first-pass bioavailability is inversely proportional to the unbound fraction, and changes in the binding are expected to have a significant effect. It is also clear that changes in both the blood flow and the intrinsic clearance of the first-pass organ may have a significant effect when the extraction ratio is high ($fu\ \text{CL}_{\text{Int}} >> Q$). On the other hand, if $Q >> fu\ \text{CL}_{\text{Int}}$ then Eq. (7) simply says that the first-pass bioavailability is approximately 1 (i.e., little or no drug is eliminated in a first pass), and changes in binding, blood flow, and intrinsic clearance are not expected to have any effect on F.

Combined Kinetic Effect of Binding Alterations

Changes in binding can occur in many situations. In some diseases, the concentrations of binding proteins may be altered, and accumulation of endogenous inhibitors of binding may occur. For example, in renal failure the albumin level may decrease, resulting in decreased binding. In renal failure, accumulation of waste products also occurs, and this, in turn, may further suppress the binding. In other diseases (e.g., inflammations), the level of α_1-acid glycoprotein increases, which may lead to increased drug binding. Concomitant administration of compounds that compete for the same binding sites may also decrease the binding and increase the unbound fraction. These changes will lead to changes in the apparent volume of distribution, in clearance, and in the first-pass availability, as discussed earlier, but will they lead to changes in the time required for the drug to reach the target organ and in the activity of the drug? The answer is not simple, as it depends on the actual values of clearance and volume of distribution, as well as on whether the target tissue is among the highly or poorly perfused tissues.

To begin answering this question we must realize that the effect of drugs in the body is related to the unbound and not the total concentration in the body [33]. Only non–protein-bound drug can interact with receptors and, therefore, elicit an effect. Therefore, it is important to look at how the ***unbound*** concentration changes. Let us assume that a situation of decreased binding exists in blood, and that no changes in the tissue binding have occurred. If a dose is given, we expect no changes in the first-pass bioavailability, if the extraction ratio in the first-pass tissue is low. If the extraction ratio is high, on the other hand, we expect a significant decrease in the first-pass bioavailability [see Eq. (8)] both of an IV administration (if first-pass in the lung is significant) or an oral administration (first-pass effect both in the liver and lung). When it comes to other routes of administration, we need to be more cautious. For example,

let us look at dermal administration. If the major first-pass elimination occurs before the drug reaches the blood (i.e., in its passage through the epithelial cells), we do not expect the blood-binding changes to be important. On the other hand, if the major first-pass elimination relates to cell downstream from where the drug enters the blood, changes in the binding will cause changes predictable from Eq. (8).

Will the drug be delayed on its way to the target tissue in the foregoing example? A decreased blood binding is equivalent to increasing the $K_{O/B}$ value. Distribution into various tissues, including the target tissue, therefore, is expected to be more extensive and swifter. However, because the drug usually has to pass the lungs before it can reach the target tissue, high $K_{O/B}$ values in the lung will significantly reduce the amount available to other tissues when the $K_{O/B}$ value increases for the lung. Consequently, a delay of distribution here can occur. On the other hand, if the $K_{O/B}$ value in the lung is low, the amount sequestered by the lung will be too small to delay the distribution. In this situation, no delay is expected in distribution of drug to the target tissue, if the target tissue is highly perfused. However, if the target tissue is a poorly perfused tissue, a delay may still occur. In this situation, a higher $K_{O/B}$ value will mean that the highly perfused tissues can take up more of the initial drug presented to them. In turn, the diffusion out of the highly perfused tissues will be slower, and may take longer before the redistribution to poorly perfused tissues is completed. Therefore, if the effect is in the poorly perfused tissues, the effect will lag.

On continuous administration, the drug's effect is dependent on the average unbound steady-state concentration. The unbound concentration is dependent on the average rate of dosing, bioavailability, clearance, and degree of binding in plasma. Let us, for the moment, assume that the bioavailability and dosing rate remain constant. The change in the steady-state unbound concentration now depends on whether the extraction ratio is low or high (or whether $Q >> fu\ CL_{Int}$ or $fu\ CL_{Int} >> Q$) in addition to the alteration of the binding. For a low-extraction ratio compound the average unbound concentration, Cu_{ss}, is

$$Cu_{ss} = \frac{F\ (\text{rate of dosing})}{CL_{Int}} \tag{9}$$

For a high-extraction ratio, the value is

$$Cu_{ss} = \frac{F\ (\text{rate of dosing})}{Q/fu} \tag{10}$$

Under the assumptions of constant bioavailability and dosing rate, a high-extraction ratio compound is expected to have an increased unbound steady-state concentration when the plasma binding is decreased, and a low-extraction ratio compound is not affected by binding changes. Now, however, we also need to evaluate whether the first-pass availability is affected by a change in the binding and adjust our expectation accordingly. Under special circumstances for which the first-pass organ is the major metabolizing organ, some simplified concepts can be established. For a low-extraction ratio compound in this situation, neither the first-pass bioavailability, nor the clearance relative to unbound drug will be affected, and the result is that there is no overall effect on the average unbound steady-state concentration and the activity of the drug. For a high-extraction ratio compound both the clearance relative to unbound drug [Eq. (5)] and the first-pass bioavailability [see Eq. (7)] will increase proportional to the increase in the unbound fraction in plasma, and the result is again that the average unbound concentration is not affected by a change in the plasma binding. If the first-pass organ is different from the major eliminating organ, the extraction ratio in these organs may be different, and their dependence on binding may differ (i.e., the unbound clearance and bioavailability may be

affected to a different extent). If saturation of the first-pass takes place, there is also likely to be a different effect of binding on the first-pass bioavailability and unbound clearance, even if the first-pass organ is the major eliminating organ.

Changes in tissue binding have effects on only the tissue distribution and not on clearance. Tissue-binding changes, therefore, are not expected to affect the first-pass bioavailability, but are expected to alter the distribution [see Eq. (3)] and the $K_{O/B}$ value. If the value of $K_{O/B}$ is reduced, we expect a smaller sequestering of drug in its passage through the body, higher initial concentrations will reach the target tissue, and the effect will occur more swiftly and, initially, more potently. But because the apparent volume of distribution is smaller, the half-life will also be smaller:

$$t_{1/2} = \frac{0.693\ V}{\mathrm{CL}} \tag{11}$$

and the concentration and the effect will fall off faster. Documentation for such changes is difficult to obtain, because we cannot measure the tissue binding directly, and we can make only inferences from overall changes in the kinetic parameters and changes in plasma-binding values [34].

A third possibility, that the binding in plasma and tissue changes to the same degree, is not expected to affect the apparent volume of distribution to the individual organs [see Eq. (3)], and the value of $K_{O/B}$ will not be changed. The clearance, on the other hand, is affected only by changes in binding in blood and will change as described in Sec. III.B, the section describing the effect of protein-binding changes on clearance. Assuming that the first-pass bioavailability is not affected, this will mean that sequestering of drug will not increase, and the total concentration reaching the target tissue should not be significantly affected. However, because the unbound fraction in plasma is increased (decreased), the unbound concentration will be increased (decreased), the unbound concentration initially reaching the target tissue will be higher (lower), and the initial pharmacological activity will be higher (lower). The duration of this change will depend on the half-life of the drug, which will be altered according to Eq. (11).

IV. SUMMARY

The extent and time course of drug action can be markedly affected by the route of drug administration into the patient as well as the pattern of drug distribution within the patient. Drug efficacy can be improved, and drug toxicity probably decreased, if the drug can be administered directly to its site of action. However, several factors prevent direct application of drugs to the site of action, including incomplete knowledge about the action site and also poor patient adherence owing to the inconvenience of using direct application formulations. Because of these factors, most drug products have been formulated as oral, solid dosage forms.

Drugs that are rapidly cleared by hepatic processes will show a decreased extent of availability following oral administration owing to metabolism of the drug on its first pass through the liver. The magnitude of this first-pass effect will depend on the blood flow to the liver and the intrinsic clearing ability of the liver (i.e., the ability of the organ to eliminate the drug independently of the rate at which drug is brought to the organ). This first-pass elimination by metabolic or biliary excretion processes can be excluded if the drug is absorbed from a sublingual site. The rectal administration of drugs eliminates approximately one-half the first-pass metabolism. Absorption through nasal, dermal, and other sites may also give rise to lower first-pass elimination and higher bioavailability than oral administration if the metabolic capacity of these sites for the drug in question is small.

Drug distribution in the patient will depend on the blood blow to various sites in the body, as well as on the partition coefficient of the drug between the blood and distributive organs. Protein binding, both in the blood and in the tissues, will markedly affect this distribution. However, free drug concentrations are generally believed to be the effective determinant in drug therapy. Often a redistribution owing to changes in protein binding will have little effect on the therapeutic efficacy, since, although total drug distribution changes, the average unbound concentrations at steady state in blood remain essentially similar. An understanding of the effects of the route of administration as well as the distribution of the drug within the body is critical to the pharmacist in planning appropriate drug dosage regimens.

REFERENCES

1. L. Z. Benet, Biopharmaceutics as a basis for the design of drug products, in *Drug Design*, Vol. 4, (E. Ariens, Ed.), Academic Press, New York, 1973, pp. 1–35.
2. L. Z. Benet, Input factors as determinants of drug activity: route, dose, dosage regimen, and the drug delivery system, in *Principles and Techniques of Human Research and Therapeutics*, Vol. 3 (F. G. McMahon, Ed.), Futura, New York, 1974, pp. 9–23.
3. T. E. Gram. The metabolism of xenobiotics by the mammalian lung, in *Extrahepatic Metabolism of Drugs and Other Foreign Compounds* (T. E. Gram, Ed.), S.P. Medical and Scientific Books, New York, 1980, pp. 159–209
4. J. R. Vane, The role of the lungs in the metabolism of vasoactive substances, in *Pharmacology and Pharmacokinetics* (T. Teorell, R. L. Dedrick, and P. G. Condliffe, Eds.), Plenum Press, New York, 1974, pp. 195–207.
5. W. L. Chiou, Potential pitfalls in the conventional pharmacokinetic studies: Effects of the initial mixing of drug in blood and the pulmonary first-pass elimination, J. Pharmacokinet. Biopharm., 7, 527–536 (1979).
6. M. Rowland, L. Z. Benet, and G. G. Graham, Clearance concepts in pharmacokinetics, J. Pharmacokinet. Biopharm., 1, 123–136 (1973).
7. E. F. Evans, J. D. Proctor, M. J. Fratkin, J. Velandia, and A. J. Wasserman, Blood flow in muscle groups and drug absorption, Clin. Pharmacol. Ther., 17, 44–47 (1975).
8. L. S. Cohen, J. E. Rosenthal, D. W. Horner, Jr., J. M. Atkins, O. A. Matthews, and S. F. Sarnoff, Plasma levels of lidocaine after intramuscular administration, Am. J. Cardiol., 29, 520–523 (1972).
9. M. L. Schwartz, M. B. Meyer, B. G. Covino, R. M. Narange, W. Sethi, A. J. Schwartz, and P. Kemp, Antiarrhythmic effectiveness of intramuscular lidocaine: Influence of different injection sites, J. Clin. Pharmacol., 14, 77–83 (1974).
10. J. T. Doluisio, J. C. LaPiana, and L. W. Dittert, Pharmacokinetics of ampicillin trihydrate, sodium ampicillin, and sodium dicloxacillin following intramuscular injection, J. Pharm. Sci., 60, 715–719 (1971).
11. A. J. Wilensky and J. A. Lowden, Inadequate serum levels after intramuscular administration of diphenylhydantoin, Neurology, 23, 318–324 (1973).
12. M. Rowland, Influence of route of administration on drug availability, J. Pharm. Sci., 61, 70–74 (1972).
13. R. N. Boyes, J. H. Adams, and B. R. Duce, Oral absorption and disposition kinetics of lidocaine hydrochloride in dogs, J. Pharmacol. Exp. Ther. 174, 1–8 (1970).
14. M. Rowland, Effect of some physiologic factors on bioavailability of oral dosage forms, in *Dosage Form Design and Bioavailability* (J. Swarbrick, Ed.), Lea & Febiger, Philadelphia, 1973, pp. 181–222.
15. D. G. Shand, E. M. Nuckolls, and J. A. Oates, Plasma propranolol levels in adults with observations in four children, Clin. Pharmacol. Ther., 11, 112–120 (1970).
16. W. H. Barr, Factors involved in the assessment of systemic or biologic availability of drug products, Drug Inf. Bull., 3, 27–45 (1969).

FRESENIUS EXHIBIT 1012
Page 127 of 293

17. M. Drucker, S. H. Blondheim, and L. Wislicki, Factors affecting acetylation in vivo of *para*-aminobenzoic acid by human subjects, Clin. Sci., 27, 133–141 (1964).
18. A. G. De Boer, D. D. Breimer, H. Mattie, J. Pronk, and J. M. Gubbens-Stibbe, Rectal bioavailability of lidocaine in man: Partial avoidance of "first-pass" metabolism, Clin. Pharmacol. Ther., 26, 701–709 (1979).
19. T. W. Schwarz, in *American Pharmacy*, 6th ed., (J. B. Sprowls, Jr. and H. M. Beal, Eds.), J. B. Lippincott, Philadelphia, 1966, pp. 311–331.
20. A. E. Pontiroli, M. Alberetto, and G. Pozza, Intranasal calcitonin and plasma calcium concentrations in normal subjects, Br. Med. J., 290, 1390–1391 (1985).
21. Y. W. Chien and K. Banga, Iontophoretic (transdermal) delivery of drugs: Overview of historic development, J. Pharm. Sci., 78, 353–354 (1989).
22. D. Parasrampuria and J. Parasrampuria, Percutaneous delivery of proteins and peptides using iontophoretic techniques, J. Clin. Pharm. Ther. 16, 7–17 (1991)
23. D. T. O'Hagan and L. Illum, Absorption of peptides and proteins from the respiratory tract and the potential for development of locally administered vaccine, Crit. Rev. Ther. Drug Carrier Syst., 7, 35–97 (1990).
24. J. J. Sciarra, Aerosols, in *Prescription Pharmacy*, 2nd ed. (J. B. Sprowls, Jr., Ed.), J. B. Lippincott, Philadelphia, 1970, pp. 280–328.
25. R. L. Dedrick and K. B. Bischoff, Pharmacokinetics in applications of the artificial kidney, Chem. Eng. Progr. Symp. Ser., 64, 32–44 (1968).
26. W. W. Mapleson, An electric analogue for uptake and exchange of inert gases and other agents, J. Appl. Physiol., 18, 197–204 (1963).
27. K. B. Bischoff and R. L. Dedrick, Thiopental pharmacokinetics, J. Pharm. Sci., 57, 1347–1357 (1968).
28. J. G. Wagner, *Fundamentals of Clinical Pharmacokinetics*. Drug Intelligence Publishers, Hamilton, IL, 1975, pp. 24–26.
29. F. Bree, G. Houin, J. Barre, J. L. Moretti, V. Wirquin, and J.-P. Tillement, Pharmacokinetics of intravenously administered ^{125}I-labelled human alpha-1-acid glycoprotein, Clin. Pharmacokinet, 11, 336–342 (1986).
30. C. G. Gahmberg and L. C. Anderson, Leucocyte surface origin of human alpha-1-acid glycoprotein (orosomucoid), J. Exp. Med., 148, 507–521 (1978).
31. T. W. Guenthert and S. Øie, Effect of plasma protein binding on quinidine kinetics in the rabbit, J. Pharmacol. Exp. Ther., 215, 165–171 (1980).
32. T. F. Blaschke, Protein binding and kinetics of drugs in liver disease, Clin. Pharmacokinet, 2, 32–44 (1977).
33. S. Øie and J.-D. Huang, Binding, should free drug levels be measured? in *Topics in Pharmaceutical Sciences 1983* (D. D. Breimer and P. Speiser, Eds.), Elsevier, Amsterdam, 1983, pp. 51–62.
34. B. Fichtl, Tissue binding of drugs: Methods of determination and pharmacokinetic consequences, in *Plasma Binding of Drugs and Its Consequences* (F. Belpaire, M. Bogaert, J. P. Tillement, and R. Verbeeck, Eds.), Academia Press, Ghent, 1991, pp. 149–158.

6

Chemical Kinetics and Drug Stability

J. Keith Guillory and Rolland I. Poust
College of Pharmacy, University of Iowa, Iowa City, Iowa

I. INTRODUCTION

In the rational design and evaluation of dosage forms for drugs, the stability of the active components must be a major criterion in determining their suitability. Several forms of instability can lead to the rejection of a drug product. First, there may be chemical degradation of the active drug, leading to a substantial lowering of the quantity of the therapeutic agent in the dosage form. Many drugs (e.g., digoxin and theophylline) have narrow therapeutic indices, and they need to be carefully titrated in individual patients so that serum levels are neither so high that they are potentially toxic, nor so low that they are ineffective. For these drugs, it is of paramount importance that the dosage form reproducibly deliver the same amount of drug.

Second, although chemical degradation of the active drug may not be extensive, a toxic product may be formed in the decomposition process. Dearborn [1] described several examples in which the products of degradation are significantly more toxic than the original therapeutic agent. Thus, the conversions of tetracycline to epianhydrotetracycline, arsphenamine to oxophenarsine, and *p*-aminosalicylic acid to *m*-aminophenol in dosage forms give rise to potentially toxic agents that, when ingested, can cause undesirable effects. Recently, Nord et al. [2] reported that the antimalarial chloroquine can produce toxic reactions that are attributable to the photochemical degradation of the substance. Phototoxicity has also been reported to occur following administration of chlordiazepoxide and nitrazepam [3]. Another example of an adverse reaction caused by a degradation product was provided by Neftel et al. [4], who showed that infusion of degraded penicillin G led to sensitization of lymphocytes and formation of antipenicilloyl antibodies.

Third, instability of a drug product can lead to a decrease in its bioavailability, rather than to loss of drug or to formation of toxic degradation products. This reduction in bioavailability can result in a substantial lowering in the therapeutic efficacy of the dosage form. This phenomenon can be caused by physical or chemical changes in the excipients in the dosage form,

FRESENIUS EXHIBIT 1012
Page 129 of 293

independent of whatever changes the active drug may have undergone. A more detailed discussion of this subject is given in Sec. II.B.

Fourth, there may be substantial changes in the physical appearance of the dosage form. Examples of these physical changes include mottling of tablets, creaming of emulsions, and caking of suspensions. Although the therapeutic efficacy of the dosage form may be unaffected by these changes, the patient will most likely lose confidence in the drug product, which then has to be rejected.

A drug product, therefore, must satisfy stability criteria chemically, toxicologically, therapeutically, and physically. Basic principles in pharmaceutical kinetics can often be applied to anticipate and quantify the undesirable changes so that they can be circumvented by stabilization techniques. Some chemical compounds, called prodrugs [5,6], are designed to undergo chemical or enzymatic conversion in vivo to pharmacologically active drugs. Prodrugs are employed to solve one or several problems presented by active drugs (e.g., short biological half-life, poor dissolution, bitter taste, inability to penetrate through the blood–brain barrier, and others). They are pharmacologically inactive as such, but are converted back in vivo to their parent (active) compounds. Naturally, the rate and extent of this conversion (which are governed by the same laws of kinetics that will be described in this chapter) are the primary determinants of the therapeutic efficacy of these agents.

In the present chapter, stability problems and chemical kinetics are introduced and surveyed. The sequence employed is as follows: first, an overview of the potential routes of degradation that drug molecules can undergo; then, a discussion of the mathematics used to quantify drug degradation; a delineation of the factors that can affect degradation rates, with an emphasis on stabilization techniques; and, finally, a description of stability-testing protocols employed in the pharmaceutical industry. It is not the intent of this chapter to document stability data of various individual drugs. Readers are referred to the compilations of stability data [7] and to literature on specific drugs [e.g., Ref. 8 and earlier volumes] for this kind of information.

II. ROUTES BY WHICH PHARMACEUTICALS DEGRADE

Since most drugs are organic molecules, it is important to recognize that many pharmaceutical degradation pathways are, in principle, similar to reactions described for organic compounds in standard organic chemistry textbooks. On the other hand, it is also important to realize that different emphases are placed on the types of reactions that are commonly encountered in the drug product stability area, as opposed to those seen in classic organic chemistry. In the latter, reactions are generally described as tools for use by the synthetic chemist; thus, the conditions under which they are carried out are likely to be somewhat drastic. Reactive agents (e.g., thionyl chloride or lithium aluminum hydride) are employed in relatively high concentrations (often > 10%) and are treated using exaggerated conditions, such as refluxing or heating in a pressure bomb. Reactions are effected in relatively short time periods (hours or days). In contrast, reactions occurring in pharmaceuticals often involve the active drug components in relatively low concentrations. For example, dexamethasone sodium phosphate, a synthetic adrenocorticoid steroid salt, is present only to the extent of about 0.4% in its injection, 0.1% in its topical cream or ophthalmic solution, and 0.05% in its ophthalmic ointment. The decomposition of a drug is likely to be mediated not by reaction with another active ingredient, but by reaction with water, oxygen, or light. Reaction conditions of interest are usually ambient or subambient. Reactions in pharmaceuticals ordinarily occur over months or years, as opposed to the hours or days required for completion of reactions in synthetic organic chemistry.

Reactions such as the Diels-Alder reaction and aldol condensations, which are important in synthetic and mechanistic organic chemistry, are of only minor importance when drug degra-

FRESENIUS EXHIBIT 1012
Page 130 of 293

dation is being considered. Students need to refocus their attention on reactions such as hydrolysis, oxidation, photolysis, racemization, and decarboxylation, the routes by which most pharmaceuticals degrade.

A cognizance of reactions of particular functional groups is important if one is to gain a broad view of drug degradation. It is a difficult task to recall degradative pathways of all commonly used drugs. Yet, through the application of functional group chemistry, it is possible to anticipate the potential mode(s) of degradation that drug molecules will likely undergo. In the following discussion, therefore, degradative routes are demonstrated by calling attention to the reactive functional groups present in drug molecules. The degradative routes are described, through the use of selected examples, as *chemical* when new chemical entities are formed as a result of drug decomposition, and as *physical* when drug loss does not produce distinctly different chemical products.

A. Chemical Degradative Routes

Solvolysis

In this type of reaction, the active drug undergoes decomposition following reaction with the solvent present. Usually, the solvent is water; but sometimes the reaction may involve pharmaceutical cosolvents, such as ethyl alcohol or polyethylene glycol. These solvents can act as nucleophiles, attacking the electropositive centers in drug molecules. The most common solvolysis reactions encountered in pharmaceuticals are those involving "labile" carbonyl compounds, such as esters, lactones, and lactams (Table 1).

Although all the functional groups cited are, in principle, subject to solvolysis, the rates at which they undergo this reaction may be vastly different. For example, the rate of hydrolysis of a β-lactam ring (a cyclized amide) is much greater than that of its linear analog. The half-life (the time needed for half the drug to decompose) of the β-lactam in potassium phenethicillin at 35°C and pH 1.5 is about 1 hr. The corresponding half-life for penicillin G is about 4 min [9]. In contrast, the half-life for hydrolysis of the simple amide propionamide in 0.18 molal H_2SO_4 at 25°C is about 58 hr [10]. It has been suggested that the antibacterial activity of β-lactam antibiotics arises from a combination of their chemical reactivity and their molecular recognition by target enzymes. One aspect of their chemical reactivity is their acylating power and, although penicillins are not very good acylating agents, they are more reactive than simple, unsubstituted amides [11]. Unactivated or "normal" amides undergo nonenzymatic hydrolysis slowly, except under the most extreme conditions of pH and temperature, because the N—C(O) linkage is inherently stable, yet when the amine function is a good leaving group (and particularly if it has a pK_a greater than 4.5), amides can be susceptible to hydrolysis at ordinary temperatures. [For a recent review on this subject see Ref. 12.] Acyl-transfer reactions in peptides, including the transfer to water (hydrolysis), are of fundamental importance in biological systems in which the reactions proceed at normal temperatures, and enzymes serve as catalysts.

The most frequently encountered hydrolysis reaction in drug instability is that of the ester, but certain esters can be stable for many years when properly formulated. Substituents can have a dramatic effect on reaction rates. For example, the *tert*-butyl ester of acetic acid is about 120 times more stable than the methyl ester, which, in turn, is approximately 60 times more stable than the vinyl analog [13]. Structure–reactivity relationships are dealt with in the discipline of physical organic chemistry. Substituent groups may exert electronic (inductive and resonance), steric, or hydrogen-bonding effects that can drastically affect the stability of compounds. Interested students are referred to a recent review by Hansch and Taft [14], and to the classic reference text written by Hammett [15].

FRESENIUS EXHIBIT 1012
Page 131 of 293

Table 1 Some Functional Groups Subject to Hydrolysis

Drug type		Examples
Esters	RCOOR'	Aspirin, alkaloids
	$ROPO_3M_x$	Dexamethasone sodium phosphate
	$ROSO_3M_x$	Estrone sulfate
	$RONO_2$	Nitroglycerin
Lactones	(R-substituted cyclic ring with C=O and O)	Pilocarpine Spironolactone
Amides	$RCONR'_2$	Thiacinamide Chloramphenicol
Lactams	(R-substituted cyclic ring with C=O and NH)	Penicillins Cephalosporins
Oximes	$R_2C = NOR$	Steroid oximes
Imides	(R-substituted cyclic ring with C=O, NH, C=O)	Glutethimide Ethosuximide
Malonic ureas	(R, R' substituted ring: C(=O)–NH, C=O, C(=O)–NH)	Barbiturates
Nitrogen mustards	$R\text{-}N(CH_2CH_2Cl)_2$	Melphalan

A dramatic decrease in ester stability can be brought about by intramolecular catalysis. This type of facilitation is affected mostly by neighboring groups capable of exhibiting acid–base properties (e.g., $-NH_2$, $-OH$, $-COOH$, and $COO-$). If neighboring-group participation leads to an enhanced reaction rate, the group is said to provide anchimeric assistance [16]. For example, the ethyl salicylate anion undergoes hydrolysis in alkaline solution at a rate that is 10^6 times larger than the experimental value for the uncatalyzed cleavage of ethyl *p*-hydroxybenzoate. The rate advantage is attributed to intramolecular general base catalysis by the phenolate anion [17].

Oxidation

Oxidation reactions are important pathways of drug decomposition. In pharmaceutical dosage forms, oxidation is usually mediated through reaction with atmospheric oxygen under ambient conditions, a process commonly referred to as autoxidation. Oxygen is, itself, a diradical, and most autoxidations are free-radical reactions. A free radical is a molecule or atom with one or more unpaired electrons. Of considerable importance to pharmaceutical scientists is a reliable method for determining and controlling oxygen concentration in aqueous solutions [18]. A thorough review of autoxidation and of antioxidants has been published [19].

The mechanisms of oxidation reactions are usually complex, involving multiple pathways for the initiation, propagation, branching, and termination steps. Many autoxidation reactions are initiated by trace amounts of impurities, such as metal ions or hydroperoxides. Thus, ferric ion catalyzes the degradation reaction and decreases the induction period for the oxidation of the compound procaterol [20]. As little as 0.0002 *M* copper ion will increase the rate of vitamin C oxidation by a factor of 10^5 [21]. Hydroperoxides contained in polyethylene glycol suppository bases have been implicated in the oxidation of codeine to codeine-*N*-oxide [22]. Peroxides apparently are responsible for the accelerated degradation of benzocaine hydrochloride in aqueous cetomacrogol solution [23] and of a corticosteroid in polyethylene glycol 300 [24,25]. Many oxidation reactions are catalyzed by acids and bases [26].

A list of some functional groups that are subject to autoxidation is shown in Table 2. The products of oxidation are usually electronically more conjugated; thus, the appearance of, or a change in, color in a dosage form is suggestive of the occurrence of oxidative degradation.

Photolysis

Normal sunlight or room light may cause substantial degradation of drug molecules. The energy from light radiation must be absorbed by the molecules to cause a photolytic reaction. If that

Table 2 Some Functional Groups Subject to Autoxidation

Functional group		Examples
Phenols	HO, R	Phenols in steroids
Catechols	HO, HO, R	Catecholamines (dopamine, isoproterenol)
Ethers	R–O–R'	Diethylether
Thiols	RCH_2SH	Dimercaprol (BAL)
Thioethers	R–S–R'	Phenothiazines (chlorpromazine)
Carboxylic acids	RCOOH	Fatty acids
Nitrites	RNO_2	Amyl nitrite
Aldehydes	RCHO	Paraldehyde

FRESENIUS EXHIBIT 1012
Page 133 of 293

energy is sufficient to achieve activation, degradation of the molecule is possible. Saturated molecules do not interact with visible or near-ultraviolet light, but molecules that contain π-electrons usually do absorb light throughout this wavelength range. Consequently, compounds such as aromatic hydrocarbons, their heterocyclic analogues, aldehydes, and ketones, are most susceptible to photolysis. In general, drugs that absorb light at wavelengths below 280 nm have the potential to undergo decomposition in sunlight, and drugs with absorption maxima greater than 400 nm have the potential for degradation both in sunlight and room light.

A dramatic example of photolysis is the photodegradation of sodium nitroprusside in aqueous solution. Sodium nitroprusside, $Na_2Fe(CN)_5NO \cdot 2H_2O$, is administered by intravenous infusion for the management of acute hypertension. If the solution is protected from light, it is stable for at least 1 year; if exposed to normal room light, it has a shelf life of only 4 hr [27].

Photolysis reactions are often associated with oxidation because the latter category of reactions can frequently be initiated by light. But, photolysis reactions are not restricted to oxidation. For sodium nitroprusside, it is believed that degradation results from loss of the nitro-ligand from the molecule, followed by electronic rearrangement and hydration. Photo-induced reactions are common in steroids [28]; an example is the formation of 2-benzoylcholestan-3-one following irradiation of cholest-2-en-3-ol benzoate. Photoadditions of water and of alcohols to the electronically excited state of steroids have also been observed [29].

Dehydration

The preferred route of degradation for prostaglandin E_2 and tetracycline is the elimination of a water molecule from their structures. The driving force for this type of covalent dehydration is the formation of a double bond that can then participate in electronic resonance with neighboring functional groups. In physical dehydration processes, such as those occurring in theophylline hydrate and ampicillin trihydrate [30], water removal does not create new bonds, but often changes the crystalline structure of the drug. Since it is possible that anhydrous compounds may have different dissolution rates compared with their hydrates [31,32], dehydration reactions involving water of crystallization may potentially affect the absorption rates of the dosage form.

Racemization

The racemization of pharmacologically active agents is of interest because enantiomers often have significantly different absorption, distribution, metabolism, and excretion, in addition to differing pharmacological actions [33]. The best-known racemization reactions of drugs are those that involve epinephrine, pilocarpine, ergotamine, and tetracycline. In these drugs, the reaction mechanism appears to involve an intermediate carbonium ion or carbanion that is stabilized electronically by the neighboring substituent group. For example, in the racemization of pilocarpine [34], a carbanion is produced and stabilized by delocalization to the enolate. In addition to the racemization reaction, pilocarpine is also degraded through hydrolysis of the lactone ring.

Most racemization reactions are catalyzed by an acid or by a base. A notable exception is the "spontaneous" racemization of the diuretic and antihypertensive agent, chlorthalidone, which undergoes facile S_N1 solvolysis of its tertiary hydroxyl group to form a planar carbonium ion. Chiral configuration is then restored by nucleophilic attack (S_N2) of a molecule of water on the carbonium ion, with subsequent elimination of a proton [35].

Incompatibilities

Chemical interactions between two or more drug components in the same dosage form, or between active ingredient and a pharmaceutical adjuvant, frequently occur. An example of

FRESENIUS EXHIBIT 1012
Page 134 of 293

drug–drug incompatibility is the inactivation of cationic aminoglycoside antibiotics, such as kanamycin and gentamicin, by anionic penicillins in IV admixtures. The formation of an inactive complex between these two classes of antibiotics occurs not only in vitro, but apparently also in vivo in patients with severe renal failure [36]. Thus, when gentamicin sulfate was given alone to patients on long-term hemodialysis, the biological half-life of gentamicin was greater than 60 hr. But, when carbenicillin disodium (CD) was given with gentamicin sulfate (GS) in the dose ratio CD/GS = 80:1, the gentamicin half-life was reduced to about 24 hr.

Many pharmaceutical incompatibilities are the result of reactions involving the amine functional group. A summary of the potential interactions that can occur between various functional groups is given in Table 3.

Other Chemical Degradation Reactions

Other chemical reactions, such as hydration, decarboxylation, or pyrolysis, also are potential routes for drug degradation. Thus, cyanocobalamin may absorb about 12% of water when exposed to air, and *p*-aminosalicylic acid decomposes with evolution of carbon dioxide to form *m*-aminophenol when subjected to temperatures above 40°C. The temperature at which pyrolytic decomposition of terfenadine occurs has been used as a criterion for determining which of several tablet excipients will be perferable for long-term stability of the drug substance [37].

B. Physical Degradative Routes

Polymorphs are different crystal forms of the same compound [38]. They are usually prepared by crystallization of the drug from different solvents under diverse conditions. Steroids, sulfonamides, and barbiturates are notorious for their propensity to form polymorphs [39]. Yang and Guillory [40] attempted to correlate the occurrence frequency of polymorphism in sulfonamides with certain aspects of chemical structure. They found that sulfonamides that did not exhibit polymorphism have somewhat higher melting points and heats of fusion than those that were polymorphic. The absence of polymorphism in sulfacetamide was attributed to the stronger hydrogen bonds formed by the amide hydrogen in this molecule. These stronger hydrogen bonds were not readily stretched or broken to form alternate crystalline structures.

Since polymorphs differ from one another in their crystal energies, the more energetic ones will seek to revert to the most stable (and the least energetic) crystal form. When several polymorphs and solvates (substances that incorporate solvent in a stoichiometric fashion into the crystal lattice) are present, the conditions under which they may interconvert can become quite complex, as is true of fluprednisolone [41].

Polymorphs may exhibit significant differences in important physicochemical parameters, such as solubility, dissolution rate, and melting point [42]. Thus, the conversion from one polymorph to another in a pharmaceutical dosage form may lead to a drastic change in the physical characteristics of the drug. A well-known example of this phenomenon is the conversion of a more soluble crystal form (form II) of cortisone acetate to a less soluble form (form V) when the drug is formulated into an aqueous suspension [43]. This phase change leads to caking of the cortisone acetate suspension.

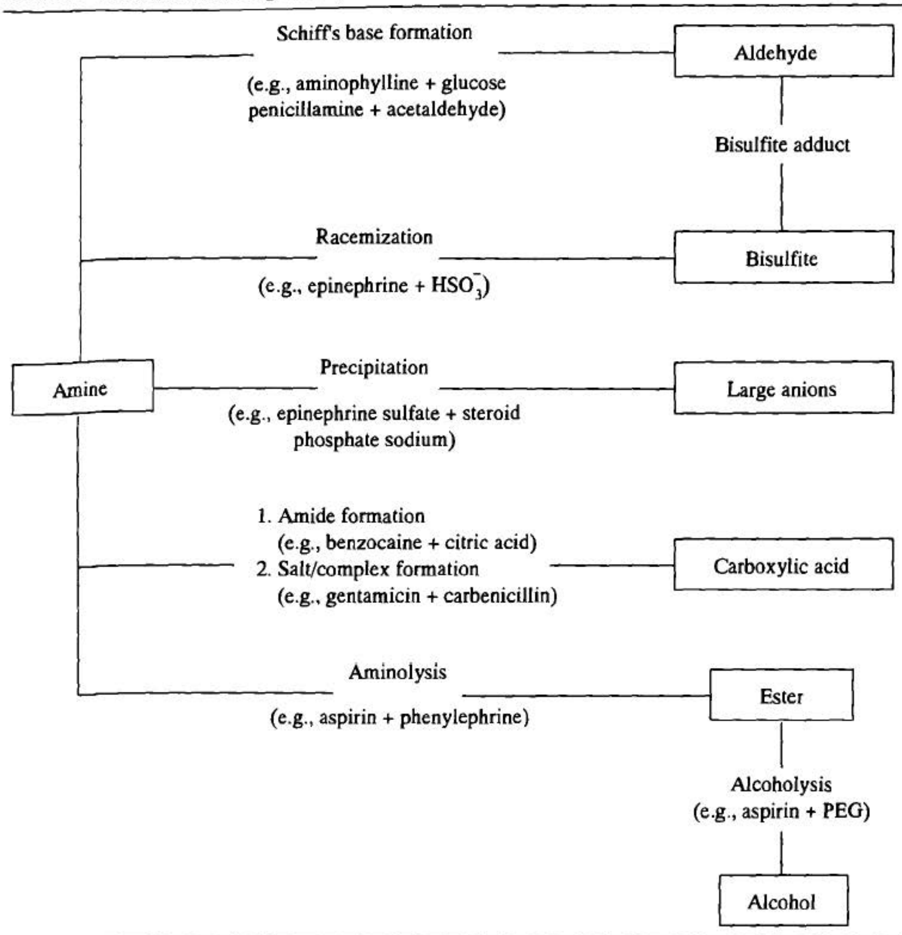
Another physical property that can affect the appearance, bioavailability, and chemical stability of pharmaceuticals is the degree of crystallinity. It has been reported that crystalline insulin [44] and crystalline cyclophosphamide [45] are much more stable than their amorphous counterparts.

Vaporization

Some drugs and pharmaceutical adjuvants possess sufficiently high vapor pressures at room temperature that their volatilization through the container constitutes a major route of drug

FRESENIUS EXHIBIT 1012
Page 135 of 293

Table 3 Some Potential Drug Incompatibilities



loss. Flavors, whose constituents are mainly ketones, aldehydes and esters, and cosolvents (low molecular weight alcohols) may be lost from the formulation in this manner. The most frequently cited example of a pharmaceutical that "degrades" by this route is nitroglycerin, which has a vapor pressure of 0.00026 mm at 20°C and 0.31 mm at 93°C [46]. Significant drug loss to the environment can occur during patient storage and use. In 1972, the Food and Drug Administration (FDA) issued special regulations governing the types of containers that may be used for dispensing sublingual nitroglycerin tablets [47].

Reduction of vapor pressure, and thereby of volatility, of drugs such as nitroglycerin can be achieved through dispersion of the volatile drug in macromolecules that can provide physicochemical interactions. The addition of macromolecules, such as polyethylene glycol, polyvinylpyrrolidone, and microcrystalline cellulose, allows preparation of "stabilized" nitroglyc-

FRESENIUS EXHIBIT 1012
Page 136 of 293

erin sublingual tablets [48,49]. A β-cyclodextrin–nitroglycerin tablet is currently being marketed in Japan to achieve the same purpose.

Another aspect of nitroglycerin instability has been observed by Fusari [49]. When conventional (unstabilized) nitroglycerin sublingual tablets are stored in enclosed glass containers, the high volatility of the drug gives rise to redistribution of nitroglycerin among the stored tablets. Interestingly, this redistribution leads to an increase in the standard deviation of the drug contents of the tablets, rather than the reverse. This migration phenomenon results in a deterioration in the uniformity of the tablets on storage.

Aging

The most interesting, and perhaps the least-reported, area of concern about the physical instability of pharmaceutical dosage forms is generally termed *aging*. This is a process through which changes in the disintegration or dissolution characteristics of the dosage form are caused by subtle, and sometimes unexplained, alterations in the physicochemical properties of the inert ingredients or the active drug in the dosage form [50]. Since the disintegration and dissolution steps may be the rate-determining steps in the absorption of a drug, changes in these processes, as a function of the "age" of the dosage form, may result in corresponding changes in the bioavailability of the drug product.

An example of this phenomenon was provided by deBlaey and Rutten-Kingma [51], who showed that the melting time of aminophylline suppositories, prepared from various bases, increased from about 20 min to over an hour after 24 weeks of storage at 22°C. Like the dissolution time for solid dosage forms, the melting time for suppositories can be viewed as an in vitro index of drug release. Thus, an increase in melting time can perceivably lead to a decrease in bioavailability. The mechanism responsible for this change appeared to involve an interaction between the ethylenediamine in aminophylline and the free fatty acids present in the suppository bases. Interestingly, no increase in melting time was detected when the suppositories were stored at 4°C, even up to 15 months.

Aging of solid dosage forms can cause a decrease in their in vitro rate of dissolution [52], but a corresponding decrease in in vivo absorption cannot be assumed automatically. For example, Chemburkar et al. [53] showed that, when a methaqualone tablet was stored at 80% relative humidity for 7–8 months, the dissolution rate, as measured by in vivo absorption, was not affected. This lack of in vitro dissolution–in vivo absorption correlation for the aged product was observed even through the particular dissolution method (that of the resin flask) was shown by the same workers to be capable of discriminating the absorption of several trial dosage forms of the same drug.

Adsorption

Drug–plastic interaction is increasingly being recognized as a major potential problem when intravenous solutions are stored in bags, or when they are infused through administration sets that are made from polyvinyl chloride (PVC). For example, up to 50% drug loss can occur after nitroglycerin is stored in PVC infusion bags for 7 days at room temperature [54]. This loss can be attributed to adsorption, rather than to chemical degradation, because the drug can be recovered from the inner surface of the container by rinsing with a less polar solvent (methanol here). A diverse array of drugs, including diazepam [55], insulin [56], isosorbide dinitrate [57], and others [58], have shown substantial adsorption to PVC. The propensity for significant adsorption is related to the oil/water partition coefficient of the drug, since this process depends on the relative affinity of the drug for the hydrophobic PVC (dielectric constant of about 3) and the hydrophilic aqueous infusion medium.

Physical Instability in Heterogeneous Systems

The stability of suspensions, emulsions, creams, and ointments is dealt with in other chapters. The unique characteristics of solid-state decomposition processes have been described in reviews by Monkhouse [59,60] and in the more recently published monograph on drug stability by Carstensen [61].

III. QUANTITATION OF RATE OF DEGRADATION

Before undertaking a discussion of the mathematics involved in the determination of reaction rates is undertaken, it is necessary to point out the importance of proper data acquisition in stability testing. Applications of rate equations and predictions are meaningful only if the data used in such processes are collected using valid statistical and analytical procedures. It is beyond the scope of this chapter to discuss the proper statistical treatments and analytical techniques that should be used in a stability study. But, some perspectives in these areas can be obtained by reading the comprehensive review by Meites [62] and from the section on statistical considerations in the stability guidelines published by the FDA in 1987 [63].

A. Kinetic Equations

Consider the reaction

$$aA + bB \rightarrow mM + nN \tag{1}$$

where A and B are the reactants; M and N, the products; and a, b, m, and n, the stoichiometric coefficients describing the reaction. The rate of change of the concentration C of any of the species can be expressed by the differential notations $-dC_A/dt$, $-dC_B/dt$, dC_M/dt, and dC_N/dt. Note that the rates of change for the reactants are preceded by a negative sign, denoting a decrease in concentration relative to time (rate of disappearance). In contrast, the differential terms for the products are positive in sign, indicating an increase in concentration of these species as time increases (rate of appearance). The rates of disappearance of A and B and the rates of appearance of M and N are interrelated by equations that take into account the stoichiometry of the reaction:

$$-\frac{1}{a}\frac{dC_A}{dt} = -\frac{1}{b}\frac{dC_B}{dt} = \frac{1}{m}\frac{dC_M}{dt} = \frac{1}{n}\frac{dC_N}{dt} \tag{2}$$

The Rate Expression

The rate expression is a mathematical description of the rate of the reaction at any time t in terms of the concentration(s) of the molecular species present at that time. By using the hypothetical reaction $aA + bB \rightarrow$ products, the rate expression can be written as

$$-\frac{dC_A}{dt} = -\frac{dC_B}{dt} \propto C^a_{A(t)}\, C^b_{B(t)} \tag{3}$$

Equation (3) in essence states that the rate of change of the concentration of A at time t is equal to that of B, and that each of these rate changes at time t is proportional to the product of the concentrations of the reactants raised to the respective powers. Note that $C_{A(t)}$ and $C_{B(t)}$ are time-dependent variables. As the reaction proceeds, both $C_{A(t)}$ and $C_{B(t)}$ will decrease in

FRESENIUS EXHIBIT 1012
Page 138 of 293

magnitude. For simplicity, these concentrations can be denoted simply by C_A and C_B, respectively.

$$-\frac{dC_A}{dt} = -\frac{dC_B}{dt} = kC_A^a C_B^b \tag{4}$$

where k is a proportionality constant, commonly referred to as the reaction rate constant or the specific rate constant. The format for rate expressions generally involves concentration terms of only the reactants and very rarely those of the products. The latter occurs only when the products participate in the reaction once it has been initiated.

The order of the reaction, n, can be defined as $n = a + b$. Extended to the general case, the order of a reaction is the numerical sum of the exponents of the concentration terms in the rate expression. Thus if $a = b = 1$, the reaction just described is said to be second-order overall, first-order relative to A, and first-order relative to B. In principle, the numerical value of a or b can be integral or fractional.

Special attention is directed to those instances in which the rate of reaction is apparently independent of the concentration of one of the reactants, even though this reactant is consumed during the reaction. For example, in the reaction between an ester and water (hydrolysis) in a predominantly aqueous environment, the theoretical rate expression for the ester can be written in terms of the concentrations of the ester (C_E) and water (C_W):

$$-\frac{dC_E}{dt} = kC_E C_W \tag{5}$$

If the initial concentration of the ester is 0.5 M or less, complete hydrolysis of the ester will bring about a corresponding decrease in the concentration of water of 0.5 M or less. Since the initial water concentration is 1000/18, which is about 55 M for an aqueous solution, the loss of water through reaction is insignificantly small and C_W can be considered a constant throughout the entire course of the reaction. Thus, in practice,

$$-\frac{dC_E}{dt} = k_\pi C_E \tag{6}$$

where $k_\pi = kC_W$. The reaction is thus apparently first-order relative to ester and zero-order relative to water; the overall reaction is known as a pseudo-first-order reaction and k_π the pseudo-first-order constant.

This type of kinetics is observed whenever the concentration of one of the reactants is maintained constant, either by a vast excess initial concentration, or by rapid replenishment of one of the reactants. Thus, if one of the reactants is the hydrogen ion or the hydroxide ion, its concentration, though probably small when compared with that of the drug, can be kept constant throughout the reaction by using buffers in the solution. Similarly, the concentration of an unstable drug in solution can be maintained invariant by preparing a drug suspension, thus providing excess solid in equilibrium with the drug in solution.

Simple Reactions

It is obvious that to quantify the rate expression, the magnitude of the rate constant k needs to be determined. Proper assignment of the reaction order and accurate determination of the rate constant is important when reaction mechanisms are to be deduced from the kinetic data. The integrated form of the reaction equation is easier to use in handling kinetic data. The integrated kinetic relationships commonly used for zero-, first-, and second-order reactions are summarized in Table 4. [The reader is advised that basic kinetic theory is also extensively exploited in pharmacokinetics; for further information on this subject, see Chapter 3.] The

Table 4 Rate Expressions for Zero-, First-, and Second-Order Reactions

	Zero-order	First-order	Second-order $a = b = c_0$	Second-order $a \neq b$
Differential rate expression	$-\frac{dc}{dt} = k$	$-\frac{dc}{dt} = kc$	$-\frac{dc}{dt} = kc^2$	$-\frac{dc}{dt} = kc_ac_b$
Integrated rate expression	$k = \frac{c_0 - c}{t}$	$k = \frac{1}{t}\ln\frac{c_0}{c}$	$\frac{1}{c} - \frac{1}{c_0} = kt$	$k = \frac{1}{t(a-b)}\ln\frac{b(a-x)}{a(b-x)}$
$t_{1/2}$	$\frac{c_0}{2k}$	$\frac{0.693}{k}$	$\frac{1}{c_0k}$	(i) When $x = 0.5a$: $\frac{1}{k(a-b)}\ln\frac{0.5ab}{a(b-0.5a)}$ (ii) When $x = 0.5b$: $\frac{1}{k(a-b)}\ln\frac{b(a-0.5b)}{0.5ab}$
$t_{90\%}$	$\frac{c_0}{10k}$	$\frac{0.105}{k}$	$\frac{0.11}{c_0k}$	(i) When $x = 0.1a$: $\frac{1}{k(a-b)}\ln\frac{0.9ab}{a(b-0.1a)}$ (ii) When $x = 0.1b$: $\frac{1}{k(a-b)}\ln\frac{b(a-0.1b)}{0.9ab}$

concentration symbols in Table 4 are defined as follows: c is the concentration of the drug at any time t and c_0 is the initial concentration. In the last column describing a second-order reaction in which the reactants A and B do not have the same initial concentrations, these are designated as a and b, respectively; x is the concentration reacted at time t.

In a reaction of either zero-order or second-order, the time to reach a certain fraction of the initial concentration, for example, $t_{1/2}$ or $t_{90\%}$ [the time required for the drug concentration to decrease to 90% of its original value (i.e. 10% degradation)] is dependent on c_0. This is illustrated in Fig. 1, in which a zero-order reaction (see Fig. 1a) and a second-order reaction (see Fig. 1b) are plotted with two initial concentrations. It is readily seen that for a zero-order reaction, the $t_{1/2}$ increases with a higher initial concentration. Conversely, for a second-order reaction, $t_{1/2}$ decreases with increasing initial concentration. For a reaction obeying first-order kinetics, the $t_{1/2}$ or $t_{90\%}$ is independent of c_0.

Complex Reactions

Parallel First-Order Reactions. In many instances, the active drug may degrade through more than one pathway:

$$A \xrightarrow{k_1} B, \quad A \xrightarrow{k_2} C, \quad A \xrightarrow{k_3} D \tag{7}$$

FRESENIUS EXHIBIT 1012
Page 140 of 293

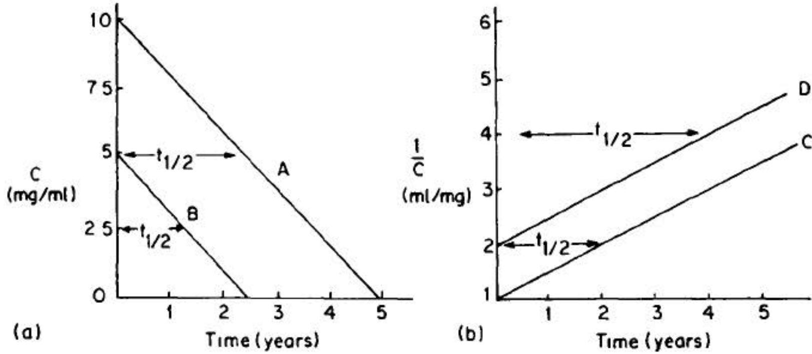


Fig. 1 Effect of initial concentration on the half-life of (a) a zero-order and (b) a second-order reaction. In (a), k = 2 mg/year-ml; curve A, initial concentration c_0 = 10 mg/ml, $t_{1/2}$ = 2.5 years; curve B, c_0 = 5 mg/ml, $t_{1/2}$ = 1.25 years. In (b), k = 0.5 ml/mg-year; curve C, c_0 = 1 mg/ml, $t_{1/2}$ = 2 years; curve D, c_0 = 0.5 mg/ml, $t_{1/2}$ = 4 years.

If the concentration of the active drug, A, can be monitored, the composite rate constant, $k' = k_1 + k_2 + k_3$, can easily be determined from the relationship $[A] = [A]_0 e^{-k't}$, where $[A]_0$ is the initial concentration and [A] is the concentration at time t. If the concentrations of A cannot be determined because of assay difficulties, it is still possible to determine k' by monitoring one of the degradation products. For example, if the concentrations of B can be assayed as a function of time, and the concentration of B at time infinity, $[B]_\infty$, is also determined, the following relationships can be derived:

$$[B] = \frac{k_1}{k'}[A]_0(1 - e^{-k't}) \tag{8}$$

$$[B]_\infty = \frac{k_1}{k'}[A]_0 \tag{9}$$

$$\ln\left(1 - \frac{[B]}{[B]_\infty}\right) = -k't \tag{10}$$

Approach to Equilibrium Through First-Order Reactions. This type of reaction can be represented by Eq. (11):

$$A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} B \tag{11}$$

The concentrations of A and B as a function of time can be derived:

$$[A] = \frac{k_2}{k_1 + k_2}[A]_0 + \frac{k_1}{k_1 + k_2}[A]_0(e^{-(k_1 + k_2)t}) \tag{12}$$

$$[B] = \frac{k_1}{k_1 + k_2}[A]_0(1 - e^{-(k_1 + k_2)t}) \tag{13}$$

The combined constants $(k_1 + k_2)$ can be obtained through Eq. (14), and the individual rate constants k_1 and k_2 can now be calculated through Eq. (15):

$$\ln([A] - [A]_\infty) = \ln[B]_\infty - (k_1 + k_2)t \tag{14}$$

$$k_1[A]_\infty = k_2[B]_\infty \tag{15}$$

Fractional Order. In the decomposition of pure solids, the kinetics of reactions can often be more complex than simple zero- or first-order processes. Carstensen [61] has reviewed the stability of solids and solid dosage forms as well as the equations that can be used in these cases. In addition to zero- and first-order kinetics, solid-state degradations are often described by fractional-order equations.

More complicated reactions schemes, including first-order reversible consecutive processes and competitive consecutive reactions, are considered in a textbook by Irwin [62]. Professor Irwin's textbook also includes computer programs written in the BASIC language. These programs can be used to fit data to the models described.

B. Energetics of Reactions

According to the transition state theory, the reaction between two molecules, A and B, to form products C and D proceeds through a transition state, X:

$$A + B \underset{}{\overset{K^\ddagger}{\rightleftharpoons}} X \longrightarrow C + D \tag{16}$$

Here $K^\ddagger$ is a thermodynamic equilibrium constant that can be expressed as a function of the activities [Eq. (17)] or of the activity coefficients γ_X, γ_A, and γ_B [Eq. (18)]:

$$K^\ddagger = \frac{a_X}{a_A a_B} \tag{17}$$

$$K^\ddagger = \frac{[X]}{[A][B]} \frac{\gamma_X}{\gamma_A \gamma_B} \tag{18}$$

The rate of the reaction $-d[A]/dt$ is proportional to the concentration of the transition state

$$-\frac{d[A]}{dt} = k'[X] \tag{19}$$

where k' is a proportionality constant. Combining Eqs. (18) and (19) yields

$$-\frac{d[A]}{dt} = k'K^\ddagger[A][B] \frac{\gamma_A \gamma_B}{\gamma_X} \tag{20}$$

If the activity coefficients are assumed to be unity, the specific rate constant k is then identical to $k'K^\ddagger$. It can be shown that

$$k' = \frac{k_B T}{h} \tag{21}$$

where k_B is Boltzmann's constant, h is Planck's constant, and T is the absolute temperature. Thus,

$$k = \frac{k_B T}{h} K^\ddagger \tag{22}$$

and

$$\ln k = \ln \frac{k_B}{h} + \ln T + \ln K^\ddagger \tag{23}$$

FRESENIUS EXHIBIT 1012
Page 142 of 293

Differentiating relative to T, we obtain

$$\frac{d \ln K^{\ddagger}}{dT} = \frac{d \ln k}{dt} - \frac{1}{T} \tag{24}$$

Since

$$\frac{d \ln K^{\ddagger}}{dT} = \frac{\Delta H^{\ddagger}}{RT^2} \tag{25}$$

where $\Delta H^{\ddagger}$ is the enthalpy of activation, Eq. (26) can be obtained by combining Eqs. (24) and (25):

$$\frac{d \ln k}{dT} = \frac{\Delta H^{\ddagger} + RT}{RT^2} \tag{26}$$

The classic Arrhenius equation is given by Eq. (27), where E_a is the energy of activation:

$$\frac{d \ln k}{dT} = \frac{E_a}{RT^2} \tag{27}$$

On comparing Eqs. (26) and (27), it follows that

$$\Delta H^{\ddagger} = E_a - RT \tag{28}$$

The other thermodynamic parameters, $\Delta G^{\ddagger}$ and $\Delta S^{\ddagger}$, the free energy and entropy of activation, respectively, can also be obtained from the foregoing relationships:

$$\Delta G^{\ddagger} = -RT \ln K^{\ddagger} = -RT \ln \frac{kh}{k_B T} \tag{29}$$

and

$$\Delta S^{\ddagger} = -\frac{\Delta G^{\ddagger} + \Delta H^{\ddagger}}{T} = R \ln \frac{kh}{k_B T} + \frac{E_a - RT}{T} \tag{30}$$

The magnitudes of the thermodynamic parameters, $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$, sometimes provide evidence supporting proposed mechanisms of drug decomposition. The enthalpy of activation is a measure of the energy barrier that must be overcome by the reacting molecules before a reaction can occur. As can be seen from Eq. (28), its numerical value is less than the Arrhenius energy of activation by the factor RT. At room temperature, RT is only about 0.6 kcal/mol. The entropy of activation can be related to the Arrhenius frequency factor (i.e., the fraction of molecules possessing the requisite energy that actually reacts). This parameter includes steric and orientation requirements of the reactants, the transition state, and the solvent molecules surrounding them. For unimolecular reactions, $\Delta S^{\ddagger}$ has a value of near zero or slightly positive. For bimolecular reactions, $\Delta S^{\ddagger}$ is more negative. For example, in the hydrolysis of esters and anhydrides, the entropy of activation is on the order of −20 to −50 entropy units, reflecting a transition state in which several solvent molecules are immobilized for solvation of the developing charges [66].

FRESENIUS EXHIBIT 1012
Page 143 of 293

IV. THE ARRHENIUS EQUATION AND ACCELERATED STABILITY TESTING

The Arrhenius equation (27) may be integrated and rewritten as Eqs. (31) and (32):

$$k = Ae^{-E_a/RT} \tag{31}$$

$$\ln \frac{k_1}{k_2} = \frac{E_a}{R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right) \tag{32}$$

where E_a is a constant and the subscripts 1 and 2 denote the two different temperature conditions. A plot of ln k as a function of $1/T$, referred to as the *Arrhenius plot*, is linear according to Eq. (31), if E_a is independent of temperature. Thus, it is possible to conduct kinetic experiments at elevated temperatures and obtain estimates of rate constants at lower temperatures by extrapolation of the Arrhenius plot. This procedure, commonly referred to as accelerated stability testing, is most useful when the reaction at ambient temperatures is too slow to be monitored conveniently and when E_a is relatively high. For example, for a reaction with an E_a of 25 kcal/mol, an increase from 25° to 45°C brings about a 14-fold increase in the reaction rate constant. In comparison, a rate increase of just threefold is obtained for the same elevation in temperature when E_a is 10 kcal/mol. The magnitude of E_a for a reaction can be obtained from the slope of its Arrhenius plot. Hydrolysis reactions typically have an E_a of 10–30 kcal/mol, whereas oxidation and photolysis reactions have smaller energies of activation [67].

An underlying assumption of the Arrhenius equation is that the reaction mechanism does not change as a function of temperature (i.e., E_a is independent of temperature). Since accelerated stability testing of pharmaceutical products normally employs a narrow range of temperature (typically, 35° to at most 70°C), it is often difficult to detect nonlinearity in the Arrhenius plot from experimental data, even though such nonlinearity is expected from the reaction mechanism [68]. Thus, even complex biological processes may show Arrhenius behavior within certain temperature ranges; Laidler [69] cited such phenomena as the frequency of flashing of fireflies and the rate of the terrapin's heartbeat as examples.

Non-Arrhenius behavior, has been observed in pharmaceutical systems [70]. This may be attributed to the possible evaporation of solvent, multiple reaction pathways, change in physical form of the formulation, and so on, [71] when the temperature of the reaction is changed. An interesting example of non-Arrhenius behavior is the increased rate of decomposition of ampicillin on freezing. Savello and Shangraw [72] showed that for a 1% sodium ampicillin solution in 5% dextrose, the percentage of degradation at 4 hr is approximately 14% at −20°C, compared with 6% at 0°C and 10% at 5°C. This decrease in stability in frozen solutions is most frequently observed when the reaction obeys second- or higher-order kinetics. For example, the formation of nitrosomorpholine from morpholine and nitrite obeys third-order kinetics [73], and the rate of nitrosation is drastically enhanced in frozen solutions (Fig. 2). A marked acceleration in the hydrolytic degradation of methyl, ethyl, and *n*-propyl 4-hydroxybenzoates in the frozen state has also been reported by Shiva et al. [74]. These authors found that although pseudo-first-order conditions found in the liquid state are also observed in the frozen state, the rate of reaction under frozen-state conditions showed very much less dependency on the initial hydroxide ion concentration.

The mechanism for rate enhancement in frozen solutions has been reviewed by Pincock [75]. In reactions following second- or higher-order kinetics, an increase in rate may be brought about by concentration of the reactants in the liquid phase, the solute molecules being excluded from the ice lattice when the solution freezes. Occasionally, an increase in rate may be due to a change in pH on freezing. Fan and Tannenbaum [73] reported that citrate-sodium hydroxide and citrate-potassium phosphate buffers do not change pH on freezing, but citrate-sodium

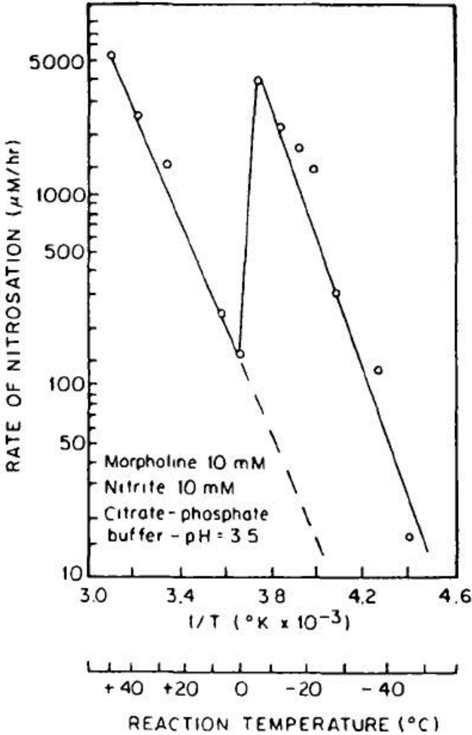


Fig. 2 Effect of temperature on the rate of nitrosation of morpholine with nitrite in citrate-sodium phosphate buffer for temperatures above and below freezing temperature. (From Ref. 73.)

phosphate buffer at pH 8 decreases to pH 3.5 and sodium hydrogen phosphate at pH 9 decreases to pH 5.5 on freezing. A possible explanation for this phenomenon is now available [76]. Monosodium phosphate forms supersaturated solutions on cooling that become amorphous, with no precipitation of the salt. The disodium and monopotassium salts, on the other hand, readily precipitated when the initial solution concentration was > 0.2 *M*. The possibility of a pH change and rate acceleration should be considered when evaluating the stability of freeze-dried products. Proteins are particularly sensitive to changes in pH, folding or unfolding to varying degrees in response to such changes. Proteins tend to be most stable at their isoelectric point owing to electrostatic interactions [77]; however, when a solution is adjusted to the optimum pH for stability at room temperature with buffers, that pH may not be maintained throughout the lyophilization cycle, and the protein may aggregate or undergo denaturation.

Considerable interest has been generated in the use of accelerated stability testing that is based on a single condition of elevated temperature and humidity. For Abbreviated New Drug Applications (ANDAs) the FDA stability guidelines [63] suggest that a tentative expiration date of 24 months may be granted for a drug product if satisfactory stability results can be documented under a stressed condition of 40°C *and* 75% relative humidity. The simplicity of such a guideline is naturally attractive because a substantial saving in time can be obtained in advancing a drug product to the marketplace. However, the reliability of prediction can be subjected to question under certain circumstances. An analysis of the use and limitation of this

approach has been presented elsewhere, and interested readers may refer to it for further information [78].

V. ENVIRONMENTAL FACTORS THAT AFFECT REACTION RATE

A rational way to develop approaches that will increase the stability of fast-degrading drugs in pharmaceutical dosage forms is through a thorough study of the factors that can affect such stability. In this section, the factors that can affect decomposition rates are discussed; it will be seen that, under certain conditions of pH, solvent, presence of additives, and so on, the stability of a drug may be drastically affected. Equations that may allow prediction of these effects on reaction rates are discussed.

A. pH

The pH of the drug solution may have a very dramatic effect on its stability. Depending on the reaction mechanism, a change of more than tenfold in rate constant may result from a shift of just 1 pH unit. When drugs are formulated in solution, it is essential to construct a pH versus rate profile so that the optimum pH for stability can be located. Many pH versus rate profiles are documented in the literature, and they have a variety of shapes. The majority of these pH versus rate profiles can be rationalized by an approach in which the reaction of each molecular species of the drug with hydrogen ion, water, and hydroxide ion is analyzed as a function of pH. The discussion that follows is divided according to the ionization capability of a drug.

1. When the drug is nonionizable: In water, three hydrolytic pathways are available [Eq. (33)], it can degrade by specific acid catalysis represented by the first kinetic term in Eq. (33), water hydrolysis (second term), and specific base catalysis (third term):

$$-\frac{dc}{dt} = k_1[H^+]c + k_2c + k_3[OH^-]c \tag{33}$$

Equation (33) can be rearranged to give Eq. (34):

$$-\frac{dc}{c\,dt} = k_{obs} = k_1[H^+] + k_2 + k_3[OH^-] \tag{34}$$

Note that k_1 and k_3 are second-order rate constants, whereas k_2 is a pseudo-first-order constant. The pH versus rate profile can be constructed by considering, in turn, that one of the three kinetic terms is predominating, thus:

(a) When $k_1[H^+] >> k_2 + k_3[OH^-]$, $k_{obs} = k_1[H^+]$ and $\log k_{obs} = \log k_1 - pH$. (35)

(b) When $k_2 >> k_1[H^+] + k_3[OH^-]$, $k_{obs} = k_2$ and $\log k_{obs} = \log k_2$ (36)

(c) When $k_3[OH^-] >> k_1[H^+] + k_2$, $k_{obs} = k_3[OH^-]$ and $\log k_{obs} = \log k_3 + pH$. (37)

Equations (35) through (37) are plotted and shown in Fig. 3a. The lines are stipled to indicate that the relative positions of the lines are not fixed, but are dependent on the relative magnitudes of the rate constants. For example, when $k_3[OH^-] > k_2 >> k_1[H^+]$, a log k_{obs} versus pH profile such as the one depicted in Fig. 3b may result. On the other hand, if $k_1[H^+]$ and $k_3[OH^-]$ are both much greater than k_2, a log k_{obs} versus pH profile may resemble the curve shown in Fig. 3c. When $k_1[H^+] > k_2 >> k_3[OH^-]$, the log k_{obs} versus pH profile will be a mirror image of Fig. 3b; and when $k_2 >> k_1[H^+] + k_3[OH^-]$, the rate constant will be pH-independent.

FRESENIUS EXHIBIT 1012
Page 146 of 293

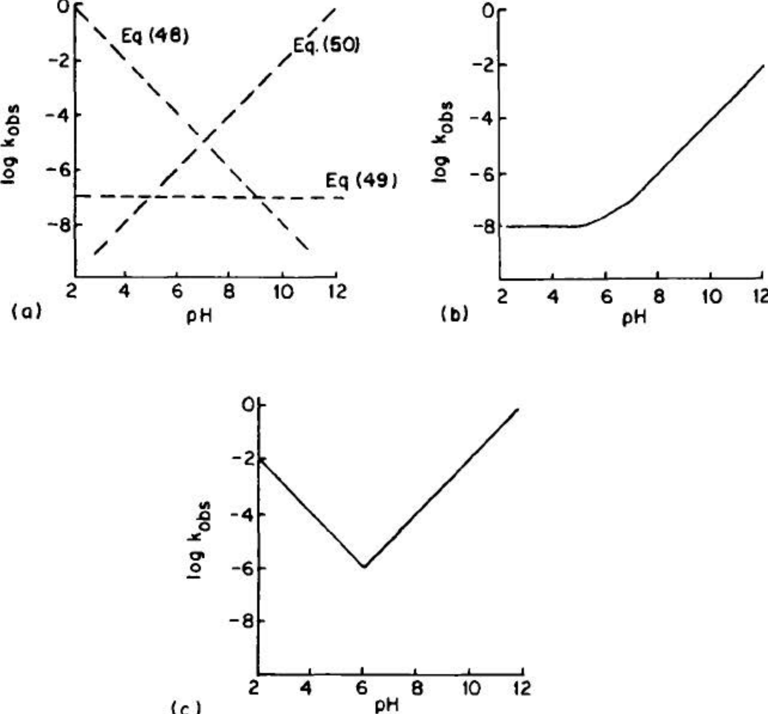


Fig. 3 The log k_{obs} versus pH profiles of nonionizable drugs.

2. When the drug is either monoacidic or monobasic, an equation similar to Eq. (34) can be written. Here, however, three kinetic terms are written for the acidic form of the drug, HA, and three terms for the basic form, A (electronic charges on HA and A are not designated here because either HA or A can be charged):

$$k_{obs} = k_1[H^+]f_{HA} + k_2 f_{HA} + k_3[OH^-]f_{HA} + k_4[H^+]f_A + k_5 f_A + k_6[OH^-]f_A \tag{38}$$

where

$$f_{HA} = \frac{[HA]}{[HA] + [A]} = \frac{[H^+]}{[H^+] + K_a} \tag{39}$$

and

$$f_A = \frac{[A]}{[HA] + [A]} = \frac{K_a}{[H^+] + K_a} \tag{40}$$

Again, Eq. (38) can be analyzed by considering each individual term as a function of pH. Since the magnitudes of both f_{HA} and f_A are dependent on the relative magnitudes of K_a and H^+, the kinetic terms can be evaluated under three conditions: (a) when $[H^+] >> K_a$, (b) when $[H^+] = K_a$, and (c) when $[H^+] << K_a$ (Table 5). The log k_{obs} versus pH profile for each kinetic term is shown in Fig. 4, using a hypothetical pK_a of 6 and the condition that $k_1 = 10^7 k_2 = k_3$

Table 5 Kinetic Expressions for Each Term in Eq. (38)

Logarithm of kinetic term	log k_{obs}		
	When $[H^+] >> K_a$	When $[H^+] = K_a$	When $K_a >> [H^+]$
$\log k_1[H^+]f_{HA}$	$\log k_1 - pH$	$\log \frac{k_1K_a}{2}$	$\log \frac{k_1}{K_a} - 2\ pH$
$\log k_2 f_{HA}$	$\log k_2$	$\log \frac{k_2}{2}$	$\log \frac{k_2}{K_a} - pH$
$\log k_3[OH^-]f_{HA}$	$\log k_3K_w + pH$	$\log \frac{k_3K_w}{2K_a}$	$\log \frac{k_3K_w}{K_a}$
$\log k_4[H^+]f_A$	$\log k_4K_a$	$\log \frac{k_4K_a}{2}$	$\log k_4 - pH$
$\log k_5 f_A$	$\log k_5K_a + pH$	$\log \frac{k_5}{2}$	$\log k_5$
$\log k_6[OH^-]f_A$	$\log k_6K_wK_a + 2\ pH$	$\log \frac{k_6K_w}{2K_a}$	$\log k_6K_w + pH$

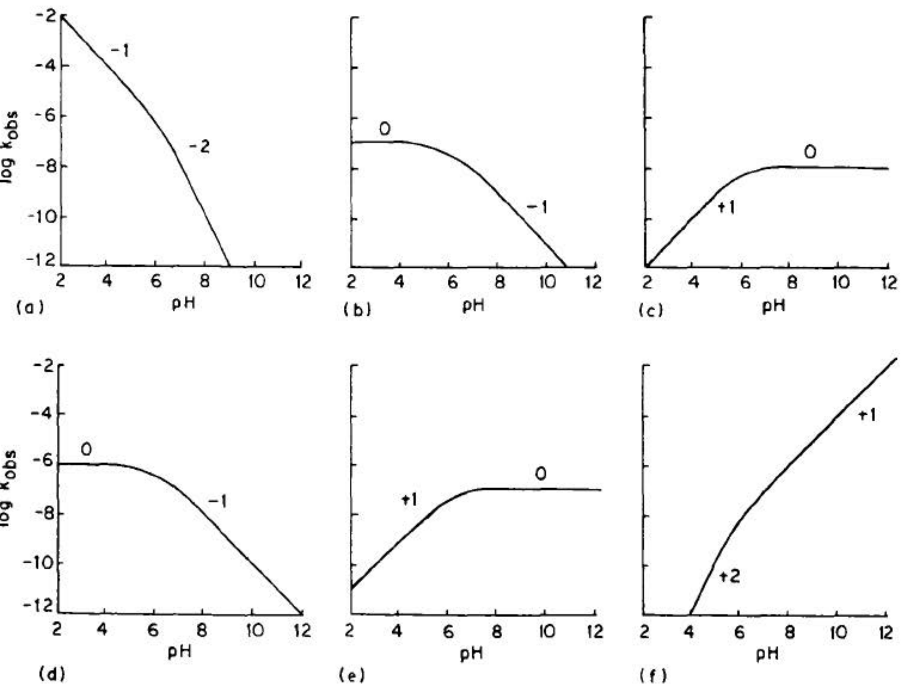


Fig. 4 The log k_{obs} versus pH profile for each kinetic term in Eq. (38): $k_1 = 10^7$, $k_2 = k_3 = k_4 = 10^7$, $k_5 = k_6 = 1$; $K_a = 10^{-6}$. Each number next to the curve indicates the slope of that portion of the curve.

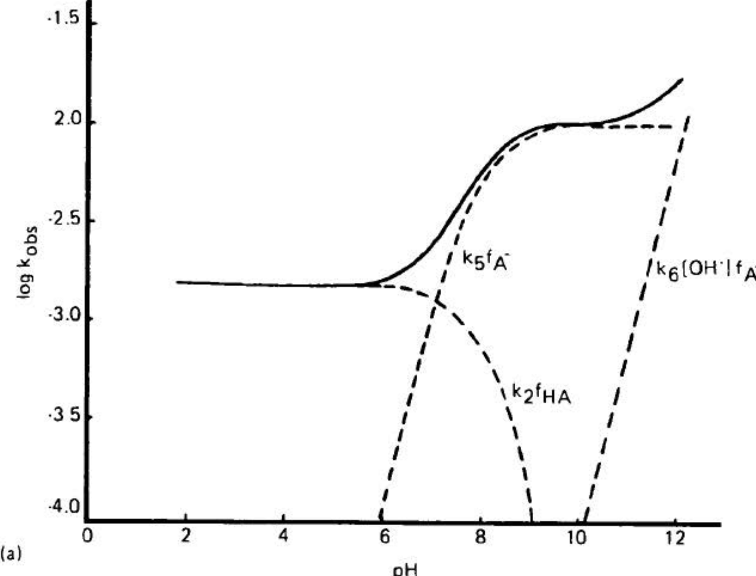


Fig. 5 (a) The pH versus log hydrolysis rate constant profile of idoxuridine at 60°C (solid line). The dashed lines indicate the individual contributions of each kinetic term. (b) pH versus log hydrolysis rate constant profile of acetylsalicylic acid (solid line). The dashed lines indicate the individual contribution of each kinetic term.

$= k_4 = 10^7\ k_5 = k_6 = 1$. Compared with the curves shown in Fig. 3a, the profiles in Fig. 4 show one break each in the lines, with a change of slope of 1 unit at the breaks. It is also seen that term (b) is equivalent to term (d) and that term (c) is equivalent to term (e), as far as their dependency on pH is concerned (see Table 5 and Fig. 4). These terms, therefore, are kinetically equivalent and are indistinguishable from each other in a rate expression. Equation (38), then, can be reduced to a combination of only four terms. The shape of the overall log k_{obs}/pH profile of any drug is determined by the relative magnitudes of the four kinetic terms over the pH range considered. Each log k_{obs} versus pH profile of a monoacidic or monobasic drug can be adequately described by a combination of no more than four terms. Figure 5 illustrates this principle by showing the log k_{obs} versus pH profiles of idoxuridine [79] and acetylsalicylic acid [80]. The hydrolysis of idoxuridine (see Fig. 5a) as a function of pH can be rationalized by the equation $k_{obs} = k_2 f_{HA} + k_5 f_{A^-} + k_6[OH^-]f_{A^-}$ (three kinetic terms), whereas the hydrolysis of acetylsalicylic acid (see Fig. 5b) can be decided only by using all four kinetic terms ($k_{obs} = k_1[H^+]f_{HA} + k_4[H^+]f_{A^-} + k_5 f_{A^-} + k_6[OH^-]f_{A^-}$.

Some compounds exhibit pH behavior in which a bell-shaped curve is obtained with maximum instability at the peak [81]. The peak corresponds to the intersection of two sigmoidal curves that are mirror images. The two inflection points imply two acid and base dissociations responsible for the reaction. For a dibasic acid (H_2A) for which the monobasic species (HA^-) is most reactive, the rate will rise with pH as $[HA^-]$ increases. The maximum rate occurs at pH = $(pK_1 + pK_2)/2$ (the mean of the two acid dissociation constants). Where an acid and base react, the two inflections arise from the two different molecules. The hydrolysis of penicillin G catalyzed by 3,6-bis(dimethylaminomethyl)catechol [82], is a typical example.

FRESENIUS EXHIBIT 1012
Page 149 of 293

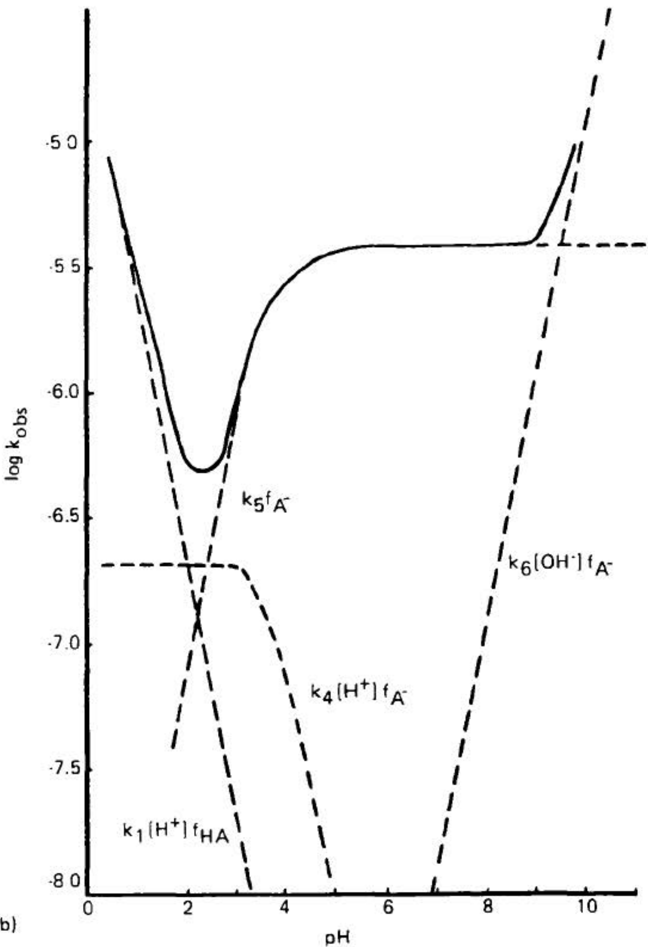


Fig. 5 Continued

B. Solvent

In many pharmaceutical dosage forms, it may be necessary to incorporate water-miscible solvents to solubilize the drug. These solvents are generally low molecular weight alcohols, such as ethanol, propylene glycol, and glycerin; or polymeric alcohols, such as the polyethylene glycols. Solvent effects can be quite complicated and difficult to predict. In addition to altering the activity coefficients of the reactant molecules and the transition state, changes of the solvent system may bring about concomitant changes in such physicochemical parameters as pK_a, surface tension, and viscosity, that indirectly affect the reaction rate. In some cases, an additional reaction pathway may be generated, or there may be a change in the product mix. The angiotensin-converting enzyme inhibitor, moexipril, undergoes hydrolysis as well as a cyclization reaction, leading to the formation of diketopiperazines. In mixed solvent (75–90% ethanol) systems the hydrolysis reaction is suppressed, but the rate of the cyclization reaction increases by 5.5-fold to 29-fold [83]. In the presence of increasing concentrations of ethanol

in the solvent, aspirin degrades by an extra route, forming the ethyl ester of acetylsalicylic acid [84]. On the other hand, a solvent change may bring about stabilization of a compound. The hydrolysis of barbiturates occurs 6.7-fold faster in water than in 50% ethanol, and 2.6-fold faster in water than in 50% glycerol [85].

Many approaches have been used to correlate solvent effects. The approach used most often is based on the electrostatic theory, the theoretical development of which has been described in detail by Amis [86]. The reaction rate is correlated with some bulk parameter of the solvent, such as the dielectric constant or its various algebraic functions. The search for empirical parameters of solvent polarity and their application in multiparameter equations has recently been intensified, and this approach is described in the book by Christian Reichardt [87]. Although the solvent effect on reaction rate could, in principle, be large, the limited availability of nontoxic solvents suitable for pharmaceutical products has rendered this stabilization approach somewhat impractical in most circumstances.

C. Solubility

As mentioned earlier in this chapter, penicillins are very unstable in aqueous solution by virtue of hydrolysis of the β-lactam ring. A successful method of stabilizing penicillins in liquid dosage forms is to prepare their insoluble salts and formulate them in suspensions. The reduced solubility of the drug in a suspension decreases the amount of drug available for hydrolysis. An example of improved stability of a suspension over that of a solution is illustrated in Fig. 6, in which a hypothetical drug is formulated as a 10-mg/ml solution (curve A) and as a suspension containing the same total amount of drug, but with a saturated solubility of 1 mg/ml (curve B). It is seen that the drug in solution undergoes first-order degradation with a half-

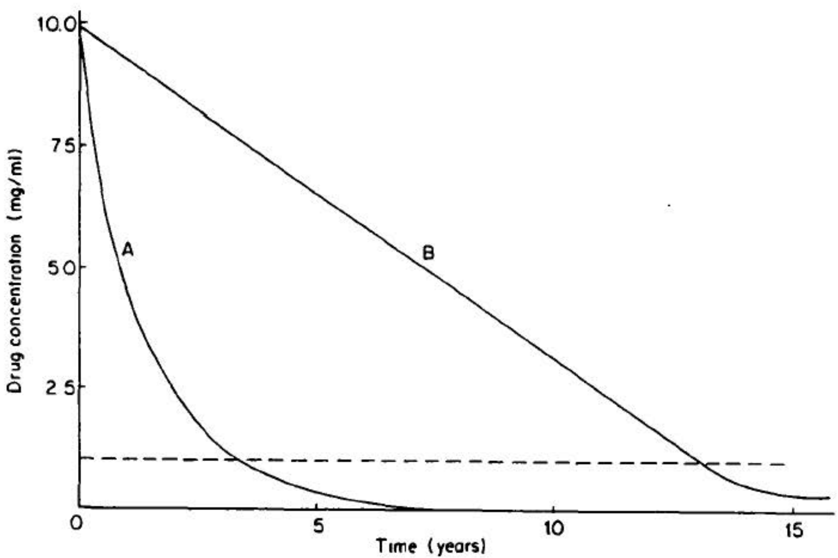


Fig. 6 Solubility effects on drug stability: curve A, drug formulated as 10 mg/ml solution ($t_{1/2}$ = 1 year); curve B, drug formulated as a suspension with a saturated solubility of 1 mg/ml ($t_{1/2}$ = 7.3 years).

life of 1 month. In the suspension, the drug degrades through zero-order kinetics until there is no more excess solid present, after which point first-order kinetics is operative.

D. Additives

Buffer Salts

In most drug solutions, it is necessary to use buffer salts to maintain the formulation at the optimum pH. These buffer salts can affect the rate of drug degradation in several ways. First, a primary salt effect results because of the effect salts have on the activity coefficient of the reactants. At relatively low ionic strengths, the rate constant, k_μ, is related to the ionic strength, μ, according to

$$\ln k_\mu = \ln k_0 + 1.02\, z_A z_B \sqrt{\mu} \tag{41}$$

where k_0 is the rate constant at $\mu = 0$; z_A and z_B are the ionic charges of reactants A and B, respectively, and the constant 1.02 is applicable to aqueous solutions at 25°C. According to Eq. (41), a plot of $\ln k_\mu$ as a function of $\sqrt{\mu}$ should yield a slope approximately equal to the product $z_A z_B$, which, in theory should always be an integer because both z_A and z_B are integral numbers. In practice, however, the slope is often fractional. The sign of the slope is sometimes informative in identifying the reactants that participate in the rate-limiting step in the reaction mechanism. But, one must avoid drawing such conclusions in instances where a choice is to be made between kinetically equivalent rate terms [88].

Buffer salts also can exert a secondary salt effect on drug stability. From Table 5 and Fig. 5, it is clear that the rate constant for an ionizable drug is dependent on its pK_a. Increasing salt concentrations, particularly from polyelectrolytes such as citrate and phosphate, can substantially affect the magnitude of the pK_a, causing a change in the rate constant. [For a review of salt effects, containing many examples from the pharmaceutical literature see Ref. 90].

Lastly, buffer salts can promote drug degradation through general acid or general base catalysis. In these cases, the rate expression will contain additional kinetic terms describing the applicable reactions between different molecular species of the drug and buffer components. The efficiency of general acid or base catalysis by the buffer components is often described by the Brønsted relationship:

$$k_A = G_A K_a^\alpha \text{ and } k_B = G_B K_B^\alpha \tag{42}$$

where k_A and k_B are the catalytic constants for general acid and base catalysis, respectively; K_a and K_a^α are the acid and base ionization constants, respectively; and G_A, G_B, α, and β are constants characteristic of the reaction, the solvent, and the temperature [91,92].

Surfactants

Addition of surface-active agents may accelerate or decelerate drug degradation. Because micellar catalysis may provide a model for enzyme reactions, acceleration of rate owing to the presence of surfactants, is well documented [93]. By comparison, stabilization of drugs through the addition of surfactants, is less frequently reported. An example in which both effects are observed is especially rare. The hydrolysis of aspirin in the plateau region (pH 6–8) is inhibited by the presence of micelles of cetyltrimethylammonium bromide and cetylpyridinium chloride, whereas in the region where the normal base-catalyzed reaction occurs (pH greater than 9) the reaction is catalyzed by micelles of these same surfactants. The mechanism of hydrolysis in the plateau region involves intramolecular general base catalysis by the adjacent ionized carboxyl group, both in the presence and absence of micelles. This reaction is inhibited in the presence of micelles because the substrate molecules are solubilized into the micelle, and water is less available in this environment than in normal aqueous solutions [94].

FRESENIUS EXHIBIT 1012
Page 152 of 293

Complexing Agents

Higuchi and Lachman [95] pioneered the work of improving drug stability by complexation. They showed that aromatic esters can be stabilized in aqueous solutions in the presence of xanthines such as caffeine. Thus, the half-lives of benzocaine, procaine hydrochloride, and tetracaine are increased by approximately two- to five-fold in the presence of 2.5% caffeine. This increase in stability is attributed to the formation of a less reactive complex between caffeine and the aromatic ester. Connors has written a comprehensive textbook that describes methods for the measurement of binding constants for complex formation in solution—along with discussions of pertinent thermodynamics, modeling statistics, and regression analysis [96]. The various experimental methods useful for measuring equilibrium constants are also discussed. A good deal of attention has recently been directed at the use of derivatives of cyclodextrin for the solubilization and stabilization of pharmaceuticals [97]. One cautionary note: complexation may adversely affect the dissolution or permeability characteristics of the drug, thereby possibly decreasing drug bioavailability.

Antioxidants and Chelating Agents

Antioxidants and chelating agents are used to protect drugs against autoxidation. Mechanistically, some antioxidants, such as ascorbic acid, ascorbyl palmitate, sodium bisulfite, sodium metabisulfite, sodium sulfite, acetone sodium bisulfite, sodium formaldehyde sulfoxylate, thioglycerol, and thioglycolic acid, act as reducing agents. They are easily oxidized, preferentially undergo autoxidation, thereby consuming oxygen and protecting the drug or excipient. They are often called oxygen scavengers because their autoxidation reaction consumes oxygen. They are particularly useful in closed systems in which the oxygen cannot be replaced once it is consumed [19]. Primary or true antioxidants act by providing electrons or labile H^+, which will be accepted by any free radical to terminate the chain reaction. In pharmaceuticals, the most commonly used primary antioxidants are butylated hydroxytoluene (BHT), butylated hydroxyanisole (BHA), the tocopherols (vitamin E), and propyl gallate. Chelating agents act by forming complexes with the heavy metal ions that are often required to initiate oxidation reactions. The chelating agents used most often are ethylenediaminetetraacetic acid (EDTA) derivatives and salts, citric acid, and tartaric acid.

E. Light and Humidity

The mathematical relationship between light intensity and drug degradation is much less developed than those describing pH and temperature effects. Part of the reason, perhaps, is that light effects on stability can be substantially avoided by using amber containers that shield off most of the ultraviolet light. Regulatory authorities usually require a statement on the photostability of products and the means of protection, if required. Often both daylight and artificial light sources are employed for tests on drug substances [98].

Humidity is a major determinant of drug product stability in solid dosage forms. Elevation of relative humidity usually decreases stability, particularly for those drugs highly sensitive to hydrolysis [99]. In addition, increased humidity can also accelerate the aging process [52,53] through interaction(s) with excipients. Humidity does not always affect drug stability adversely. Cyclophosphamide, in lyophilized cakes containing mannitol or sodium bicarbonate, undergoes rapid ($t_{90} \approx 15$ days) degradation in the solid state. The cyclophosphamide was in the amorphous state in these formulations. However, on exposure to high humidity, the cyclophosphamide was converted to the crystalline monohydrate form, which exhibited greatly improved stability [45]. Reviews dealing with the effects of moisture on the physical and chemical stability of drugs are available [100,101]. As peptides and proteins have become more impor-

tant as therapeutic agents, the role residual moisture plays in their stabilization has attracted a good deal of attention [102,103].

VI. STABILITY TESTING IN THE PHARMACEUTICAL INDUSTRY

Stability testing of drug substances and drug products begins as part of the drug discovery and synthesis or development–preformulation effort and ends only with the demise of the compound or commercial product. Activities include testing of drug substance, of compatibility with excipients, of preclinical formulations, of Phase I formulations and modifications, of the final, NDA (commercial) formulation, and of postapproval formulation changes. The regulatory basis for the various aspects of stability testing is established in 21 CFR 211.137, 211.160, 211.170, 211.190, 314.50, 314.70, and 314.81 [104–110]. In addition, FDA guidelines for submitting stability data were published several years ago [63].

The final draft of the ICH Harmonized Tripartite Guideline "Stability Testing of New Drug Substances and Products" was issued by the International Conference on Harmonization (ICH) Expert Working Group of the ICH on technical requirements for the registration of pharmaceuticals for human use, in October 1993 [111]. These guidelines provide definitions of key terms and principles used in the stability testing of drug substance and drug product and cover a scope of issues similar to those appearing in the FDA guideline [63]. The draft has been recommended for adoption to the regulatory bodies of the participating parties.

Another valuable source of information on stability testing can be found in Carstensen's recent book [61].

A. Resources

Personnel

The number and types of personnel are dictated by the size of the program, the functions contained within the program, and the nature of the program. Some companies maintain separate development and commercial product programs; other integrate the two. It is important, both for operating efficiency and regulatory compliance, that all personnel, regardless of their function, receive adequate and well-documented training in both current Good Manufacturing Practices (cGMPs) and in the technical aspects of their jobs.

Education and Experience. The program is generally headed by a professional with several years experience in the company. Experience in some aspect of formulation development or in the stability-testing program itself may be of equal importance to the educational level of the person heading the program. Successful programs have been led by people at the bachelors level with a degree in pharmacy, chemistry, or related science, as well as by people with masters or doctoral level degrees. If the size of the program warrants, intermediate-level scientists, usually at the bachelors level, may assume responsibility for specific functions within the program (e.g., chemical and physical testing, documentation, and so on). Technicians generally have a high-school education or equivalent with clerical or scientific experience or interest. Clerical and data-entry personnel have traditional training and experience in their respective areas.

Organization. Functions are often divided into chemical and physical testing, documentation, and clerical–computer operations, if the information system is computerized. Successful manual or paper systems are possible without the aid of computers. However, custom-designed software or commercially available database programs can also be programmed to automate the program. The documentation function usually consists of one or more persons who prepare the stability sections of regulatory documents. Persons specifically trained in technical writing

FRESENIUS EXHIBIT 1012
Page 154 of 293

or scientists with an interest and talent in document preparation generally perform well in this capacity.

Facilities

Storage Chambers. Chambers capable of accurately maintaining freezer conditions (−20° to −10°C), refrigerator (2°–8°C), and controlled room temperature (15°–30°C) are a necessity. Also, elevated temperature (40°C and 50°C) and humidity (40°C/75% RH) conditions should be available. Finally, a high-intensity light cabinet and a cycling chamber capable of cycling both temperature and humidity are needed. These chambers should be capable of controlling temperature within ± 2°C and humidity within ± 5%. They should be calibrated periodically according to a standard operating procedure and records of these calibrations maintained in a log book for each chamber.

The ICH Tripartite Guidelines have established that long-term stability testing should be done at 25°C/60% RH. Stress testing should be done at 40°C/75% RH for 6 months. If "significant change" occurs at these stress conditions, then the formulation should be tested at an intermediate condition (i.e., 30°C/75% RH). *Significant change* is defined in the guidelines.

Storage chambers should be validated for their ability to maintain the desired conditions and, if so equipped, the ability to sound an alarm if a mechanical or electrical failure causes the temperature to deviate from preestablished limits. They should also be equipped with recording devices that will provide a continuous and permanent history of their operation. Log books should be maintained and frequent readings of mercury-in-glass, National Institute of Science and Technology-traceable thermometers recorded.

Bench Space. Adequate laboratory bench, desk, and file space are needed for physical, chemical, and microbiological testing, for documentation, and for storing records, respectively.

Equipment

Chemical Testing. Adequate instrumentation for a variety of different test methods should be available. Most stability-indicating chemical assays are performed by high-performance liquid chromatography (HPLC). Occasionally, gas chromatography, infrared spectrophotometry, or spectrofluorimetry are used. Test methods should be validated [112–114] and stability-indicating (i.e., able to distinguish the active ingredient from its degradation products) so that the active can be accurately measured. Also, methods are needed for identifying and quantitating degradation products that are present at levels of 0.1% or greater.

Biological Testing. A portion of the laboratory may be reserved for biological testing, or this work can be done by the company's microbiological laboratory. The ability to perform sterility, pyrogen, limulus amebocyte lysate (LAL), preservative challenge, and bioburden tests is needed to support the stability program. As in the chemical assays, test methods should be validated and operator familiarity should be documented.

Physical Testing. Equipment and trained personnel should be available for performing such tests as pH, tablet hardness, etc. One important and sometimes overlooked aspect of physical testing is the recording of product appearance. Carefully defined descriptions of appearance and standard descriptions of changes in appearance should be developed, especially when there is a high probability that the person who made the observation at the previous sampling time will not be the person making the observation at the next sampling time. Some companies maintain samples at a lower-than-label storage condition (e.g., refrigeration) to use as standards, assuming that minimal or no appearance change will occur at this condition. The same argument for standard nomenclature applies to other test parameters that are subjective in nature.

FRESENIUS EXHIBIT 1012
Page 155 of 293

Computers. A certain number of personal computers are necessary for report generation and regulatory submission preparation. In addition, these may be useful for record keeping, depending on the type of stability information system that the company chooses to use. Alternatively, if the information system is intended to be accessible (read only) to many users, it may be more efficient to develop a mainframe system, rather than a local area network of minicomputers. The size of the database will help determine the nature of the software and hardware configuration used for this function.

B. Program

Scope and Goals

Activities encompassed by the stability program include sample storage of either development or production batches (or both); data collection and storage–retrieval; physical, chemical, and microbiological testing; document preparation of regulatory submissions; and package evaluation. In certain companies, some of these functions (e.g., regulatory document preparation), may be performed by personnel in separate departments. Nonetheless, the function is part of the company's overall stability program.

Protocols

The FDA stability guidelines [63] and ICH Harmonized Tripartite Guidelines [111] are rather detailed concerning sampling times, storage conditions, and specific test parameters for each dosage form. Generally, samples stored at the label storage condition, controlled room temperature for most products, are tested initially and after 3, 6, 9, 12, 18, and 24 months; and annually, thereafter. Accelerated testing is generally done more frequently and for a shorter duration (e.g., 1, 2, 3, and 6 months). Three batches should be tested to demonstrate batch-to-batch uniformity. The number three represents a compromise between a large number desired for statistical precision and the economics of maintaining a manageable program. Generally, real-time data obtained at the label storage conditions on the final formulation in the final packaging configuration(s) are needed for an NDA. Supportive data obtained from drug substance stability studies, preformulation studies, and investigational formulations tested during clinical trials and formulation development may be used to supplement primary stability data. Requirements for the IND are less defined, the only requirement being that there should be adequate data to support the clinical batch(es) for the duration of the trials.

There are instances, especially with solid, oral dosage forms, for which several package types and configurations are desired by marketing, and three or more strengths are needed for flexibility in dosing. In these situations, it may be feasible to apply the principles of bracketing or matrixing to reduce the amount of testing. Bracketing refers to reduced testing of either an intermediate dosage strength or package size when the formulation characteristics of all strengths are virtually identical, or when the same container–closure materials are used for all package sizes. Matrixing refers to reduced testing, regardless of strength or container, in situations for which there are similarities in formulation or container–closure. Bracketing and matrixing are acceptable only when the product is chemically and physically very stable and does not interact with the container–closure. Demonstration of this chemical and physical stability must be documented by preformulation, drug substance stability, and early formulation stability data. Although not as common, it may be possible to employ bracketing and matrixing with other types of dosage forms. It is imperative that discussions of such strategies with FDA should always occur before implementation.

Documentation

The need for adequate documentation of laboratory operations is established not only by good science, but also by regulatory requirements [107].

FRESENIUS EXHIBIT 1012
Page 156 of 293

Documentation of all facets of the operation is necesary. This includes validation and periodic calibration of storage chambers, instrumentation, and computer programs. Log books for the storage chambers and instruments are also necessary. Standard operating procedures are needed for, among other things, the stability program itself, use of instrumentation, documentation of experiments and their results, determination of expiration dates, investigation of specification failures, and operation of a computerized record-keeping system.

Many companies have developed or purchased computer software for the purpose of storing stability data for a large number of studies. One example of a commercially available system is called "Stability System" [115]. This system can perform other functions as well, including work scheduling, preparation of summaries of selected or all studies in the system, tabulation of data for individual studies, label printing, statistical analysis and plotting, and search capabilities. Such systems should be validated to keep pace with current regulatory activity [116].

C. Regulatory Concerns

Current Good Manufacturing Practice Compliance

Current Good Manufacturing practices [105] establish the requirements for maintaining a stability program and require that most pharmaceutical dosage forms have an established expiration date, supported by test data [104]. There are few allowable exceptions.

Food and Drug Administration Stability Guidelines

The guidelines under which stability programs operate and corresponding documentation is prepared were issued in 1987 [63]. Revisions of these guidelines are currently in preparation. Although the agency emphasizes that these are guidelines, and not regulations, it is generally prudent to follow specific recommendations as indicated in the guidelines. Deviations or omissions should be addressed, and the reasons should be supported with data when applicable.

Regulatory Submissions

An easy-to-read stability summary document will go a long way toward rapid approval of any regulatory submission. Such a document should include a number of items. A clear statement of the objective(s) of the studies included in the submission and the approach that was taken to achieve the objective(s) is critical. This statement of objective(s) should accompany basic information including product and drug substance names, dosage forms and strengths, and type(s) of container–closure systems. Although the objective is usually stated in the summary letter accompanying the submission, a brief reminder to the reviewing chemist is helpful.

A discussion of each of the parameters that were tested in the course of the evaluation, including test methods and specifications for each, should then follow. These parameters should follow those recommended in the stability guidelines [63] for the specific dosage form. It is especially important to provide a rationale for those parameters not studied. Next, should come the study design itself, which should include a list of batch identification number, size, and date of manufacture, as well as packaging configuration, storage conditions, and sampling times for each batch. The strategy and rationale for any bracketing or matrixing should also be presented.

The actual data, including replicates, mean, and range, in tabular form should follow accompanied by a brief discussion of the data. It is important to explain any out-of-specification data. Statistical analyses for all parameters that lend themselves to such analyses, along with conclusions, should be incorporated into the document at this point. These statistical analyses should be accompanied by the results of experiments conducted to determine the "poolability" of batches, or commonality of slopes and intercepts of individual batches. Graphs of these data should be included as part of the documentation.

FRESENIUS EXHIBIT 1012
Page 157 of 293

Protocols for these batches and a commitment to continue them along with a "tentative" expiry date should also be included. Approval of these protocols will allow extension of the expiry date without a special supplement as long as the data remain within specifications. These data will ultimately be reported to FDA as part of periodic reports following NDA approval. Protocols intended for use on commercial batches should also be submitted.

Finally, the three-part commitment to mount studies for the first three production batches and a statistically determined number (at least one) each year, to update current studies in annual reports, and to withdraw any lots not meeting specifications should appear in the submission. Statistical sampling of production batches is usually based on $\log N$, $\sqrt{N}$, $\cdots$, where N is the number of batches produced per year. These batches are generally spread over various package types and manufacturing campaigns. There should be a standard operating procedure to handle specification deviations, including confirmation of the results, cause-and-effect investigation, impact analysis, final report to management, and field alert or batch recall notice to FDA.

Annual Product Review

Once a product gains FDA approval for marketing, the sponsor should maintain a readily retrievable profile of commercial batches. This includes individual batch release data and stability data. These data should be compiled throughout the year and tabulated before the anniversary of NDA approval for submission in the annual product report to FDA. By maintaining an ongoing database that is reviewed as new information is added, changing trends in the data can be observed and management notified if any of these trends are unfavorable.

REFERENCES

1. E. H. Dearborn, in *The Dating of Pharmaceuticals* (J. J. Windheuser and W. L. Blockstein, eds.), University Extension, University of Wisconsin, Madison, WI, 1970, p. 29.
2. K. Nord, J. Karlsen, and H. H. Tønnesen, Int. J. Pharm., 72, 11 (1991).
3. P. J. G. Cornelissen, G. M. J. Beijersbergen van Henegouwen, and K. W. Gerritsma, Int. J. Pharm., 1, 173 (1978).
4. K. A. Neftel, M. Walti, H. Spengler, and A. L. deWeck, Lancet, 1, 986 (1982).
5. V. J. Stella, T. J. Mikkelson, and J. D. Pipkin, in *Drug Delivery Systems* (R. L. Juliano, ed.), Oxford University Press, New York, 1980.
6. A. A. Sinkula, in *Sustained and Controlled Release Drug Delivery Systems* (J. R. Robinson, ed.), Marcel Dekker, New York, 1978.
7. K. A. Connors, G. L. Amidon, and V. J. Stella, *Chemical Stability of Pharmaceuticals: A Handbook for Pharmacists*, 2nd ed., John Wiley & Sons, New York, 1986.
8. H. G. Brittain, ed., *Analytical Profiles of Drug Substances and Excipients*, Vol. 21, Academic Press, San Diego, 1992.
9. M. A. Schwartz, A. P. Granatek, and F. H. Buckwalter, J. Pharm. Sci., 51, 523 (1962).
10. V. K. Krieble and K. A. Holst, J. Am. Chem. Soc., 60, 2976 (1938).
11. L. A. Casey, R. Galt, and M. I. Page, J. Chem. Soc. Perkin Trans., 2, 23 (1993).
12. R. S. Brown, A. J. Bennet, and H. Slebocka-Tilk, Acc. Chem. Res., 25, 481 (1992).
13. H. B. Mark, Jr. and G. A. Rechnitz, in *Chemical Analysis*, Vol. 24 (P. J. Elving and I. M. Kolthoff, eds.), Wiley-Interscience, New York, 1970.
14. C. Hansch, A. Leo, and R. W. Taft, Chem. Rev., 91, 165 (1991).
15. L. P. Hammett, *Physical Organic Chemistry*, 2nd Ed., McGraw-Hill, New York, 1970.
16. B. Capon and S. P. McManus, *Neighboring Group Participation*, Plenum Press, New York, 1976.
17. M. N. Khan and S. K. Gambo, Int. J. Chem. Kinetics, 17, 419–428 (1985).
18. R. E. Lindstrom, S. N. Patel, and P. K. Wilkerson, J. Parenter. Drug Assoc., 34, 5 (1980).

FRESENIUS EXHIBIT 1012
Page 158 of 293

19. D. M. Johnson and L. C. Gu, Autoxidation and antioxidants, in *Encyclopedia of Pharmaceutical Technology* (J. Swarbrick and J. C. Boylan, eds.), Marcel Dekker, New York, 1988, pp. 415–449.
20. T. M. Chen and L. Chafetz, J. Pharm. Sci., 76, 703 (1987).
21. P. Finholt, H. Kristiansen, L. Kyowezynski, and T. Higuchi, J. Pharm. Sci., 55, 1435 (1966).
22. J. Schulz and K.-H. Bauer, Acta Pharm. Technol., 32, 78 (1986).
23. R. Hamburger, E. Azaz, and M. Donbrow, Pharm. Acta Helv., 50, 10 (1975).
24. J. W. McGinity, J. A. Hill, and A. L. La Via, J. Pharm. Sci. 64, 356 (1975).
25. J. W. McGinity, T. R. Patel, A. H. Naqvi, and J. A. Hill, Drug Dev. Commun., 2, 505 (1976).
26. L. Gu, H.-S. Chiang, and D. M. Johnson, Int. J. Pharm., 41, 105 (1988).
27. M. J. Frank, J. B. Johnson, and S. H. Rubin, J. Pharm. Sci., 65, 44 (1976).
28. K. Thoma and R. Kerker, Pharm. Ind., 54, 551 (1992).
29. J. A. Waters, Y. Kondo, and B. Witkop, J. Pharm. Sci., 61, 321 (1972).
30. E. Shefter, H.-L. Fung, and O. Mok, J. Pharm. Sci., 62, 791 (1973).
31. E. Shefter and T. Higuchi, J. Pharm. Sci., 52, 781 (1963).
32. S. R. Byrn, *Solid-State Chemistry of Drugs*, Academic Press, New York, 1982.
33. F. Jamali, R. Mehon, F. M. Pasutto, J. Pharm. Sci., 78, 695 (1989).
34. M. A Nunes and E. Brochmann-Hanssen, J. Pharm. Sci., 63, 716 (1974).
35. G. Severin, Chirality, 4, 111–116 (1992).
36. L. J. Riff and G. G. Jackson, Arch. Intern. Med., 130, 887 (1972).
37. M. D. Santos-Buelga, M. J. Sanchez-Martin, and M. Sanchez-Camazano, Thermochim. Acta, 210, 255 (1992).
38. J. Haleblian and W. McCrone, J. Pharm. Sci., 58, 911 (1969).
39. M. Kuhnert-Brandstätter, *Thermomicroscopy in the Analysis of Pharmaceuticals*, Pergamon Press, Oxford, 1971, pp. 37–42.
40. S. S. Yang and J. K. Guillory, J. Pharm. Sci., 61, 26 (1972).
41. J. Haleblian, R. T. Koda, and J. A. Biles, J. Pharm. Sci., 60, 1485 (1971).
42. D. J. W. Grant and T. Higuchi, *Solubility Behavior of Organic Compounds*, John Wiley & Sons, New York, 1990.
43. T. J. Macek, U.S. Patent 2,671,750, March 9, 1954.
44. J. Brange and L. Langkjaer, Acta Pharm. Nord., 4(3), 149–158 (1992).
45. T. R. Kovalcik and J. K. Guillory, J. Parenter. Sci. Technol., 42, 29 (1988).
46. S. Budavari, ed., *The Merck Index*, 11th Ed., Merck & Co., Rahway, NJ, 1989.
47. Fed. Regist., 37, 15959 (1972).
48. H.-L. Fung, S. K. Yap, and C. T. Rhodes, J. Pharm. Sci., 63, 1810 (1974).
49. S. A. Fusari, J. Pharm. Sci., 62, 2021 (1973).
50. Z. Chowhan, Pharm. Technol. 6(9), 47–65 (1982).
51. C. J. deBlaey and J. J. Rutten-Kingma, Pharm. Acta Helv., 51, 186 (1976).
52. S. T. Horhota, J. Burgio, L. Lonski, and C. T. Rhodes, J. Pharm. Sci., 65, 1746 (1976).
53. P. B. Chemburkar, R. D. Smyth, J. D. Buehler, P. B. Shah, R. S. Joslin, A. Polk, and N. H. Reavey-Cantwell, J. Pharm. Sci., 65, 529 (1976).
54. B. L. McNiff, E. F. McNiff, and H.-L. Fung, Am. J. Hosp. Pharm., 36, 173 (1979).
55. W. A. Parker, M. E. Morris, and C. A. Shearer, Am. J. Hosp. Pharm., 36, 505 (1979).
56. J. I. Hirsch, J. H. Wood, and R. B. Thomas, Am. J. Hosp. Pharm., 38, 995 (1981).
57. P. A. Cossum and M. S. Roberts, Eur. J. Clin. Pharmacol., 19, 181 (1981).
58. E. A. Kowaluk, M. S. Roberts, H. D. Blackburn, and A. E. Pollack, Am. J. Hosp. Pharm., 38, 1308 (1981).
59. D. C. Monkhouse and L. Van Campen, Drug Dev. Ind. Pharm., 10, 1175 (1984).
60. D. C. Monkhouse, Drug Dev. Ind. Pharm., 10, 1373 (1984).
61. J. T. Carstensen, *Drug Stability, Principles and Practices*, Marcel Dekker, New York, 1990.
62. L. Meites, CRC Crit. Rev. Anal. Chem., 8, 55 (1979).
63. Guidelines for Submitting Documentation for the Stability of Human Drugs and Biologics, February, 1987, Center for Drugs and Biologics, Food and Drug Administration, Rockville, MD.
64. J. T. Carstensen, J. Pharm. Sci., 63, 1 (1974).

65. W. J. Irwin, *Kinetics of Drug Decomposition, Basic Computer Solutions*, Elsevier Science, Amsterdam, 1990.
66. W. P. Jencks, *Catalysis in Chemistry and Enzymology*, McGraw-Hill, New York, 1969.
67. E. R. Garrett, Adv. Pharm. Sci., 2, 1 (1976).
68. H.-L. Fung and S.-Y. P. King, in *Pharm Tech Conference '83 Proceedings*, Aster Publishing, Springfield, OR, 1983.
69. K. J. Laidler, J. Chem. Educ., 49, 343 (1972).
70. M. J. Pikal, A. L. Lukes, and J. E. Lang, J. Pharm. Sci., 66, 1312 (1977).
71. A. J. Woolfe and H. E. C. Worthington, Drug Dev. Commun., 1, 185 (1974).
72. D. R. Savello and R. F. Shangraw, Am. J. Hosp. Pharm., 28, 754 (1971).
73. T.-Y. Fan and S. R. Tannenbaum, J. Agric. Food Chem., 21, 967 (1973).
74. R. Shija, V. B. Sunderland, and C. McDonald, Int. J. Pharm., 80, 203 (1992).
75. R. E. Pincock, Acc. Chem. Res., 2, 97 (1969).
76. N. Murase, P. Echlin, and F. Franks, Cryobiology, 28, 364–375 (1991).
77. T. Chen, Drug Dev. Ind. Pharm., 18, 1311 (1992).
78. H.-L. Fung and S.-Y. P. King, in *Pharm Tech Conference '83 Proceedings*, Aster Publishing, Springfield, OR, 1983.
79. L. J. Ravin, C. A. Simpson, A. F. Zappala, and J. J. Gulesich, J. Pharm. Sci., 53, 106 (1964).
80. E. R. Garrett, J. Am. Chem. Soc., 79, 3401 (1957).
81. J. I. Wells, *Pharmaceutical Preformulation*, Ellis Horwood, West Sussex, UK, 1988.
82. M. A. Schwartz, J. Pharm. Sci., 53, 1433 (1964).
83. L. Gu and R. G. Strickley, Int. J. Pharm., 60, 99 (1990).
84. E. R. Garrett, J. Org. Chem., 26, 3660 (1961).
85. K. Thoma and M. Stuve, Pharm. Ind., 47, 1078–1081 (1985).
86. F. S. Amis, *Solvent Effects on Reaction Rates and Mechanisms*, Academic Press, New York, 1966.
87. C. Reichardt, *Solvents and Solvent Effects in Organic Chemistry*, 2nd Ed., VCH Verlagsgesellschaft, Weinheim, 1988.
88. K. A. Connors, *Chemical Kinetics: The Study of Reaction Rates in Solution*, VCH Publishers, New York, 1990, p. 411.
89. J. W. Moore and R. G. Pearson, *Kinetics and Mechanism*, Wiley-Interscience, New York, 1981.
90. J. T Carstensen, J. Pharm. Sci., 59, 1140 (1970).
91. B. G. Cox, A. J. Kresge, and P. E. Sørensen, Acta Chem. Scand., A24, 202 (1988).
92. A. J. Kresge, Chem. Soc. Rev., 2, 475 (1973).
93. C. A. Bunton and G. Savelli, in *Advances in Physical Organic Chemistry*, Vol. 22 (V. Gold and D. Bethell, eds.), Academic Press, Orlando, FL, 1986, pp. 213–309.
94. T. J. Broxton, Aust. J. Chem., 35, 1357 (1982).
95. T. Higuchi and L. Lachman, J. Am. Pharm. Assoc. (Sci. Ed.), 44, 521 (1955).
96. K. A. Connors, *Binding Constants: The Measurement of Molecular Complex Stability*, John Wiley & Sons, New York, 1987.
97. O. Bekers, E. V. Uijtendaal, J. H. Beijnen, A. Bult, and W. J. M. Underberg, Drug Dev. Ind. Pharm., 17, 1503 (1991).
98. N. H. Anderson, D. Johnson, M. A. McLelland, and P. Munden, J. Pharm. Biomed. Anal., 9, 443 (1991).
99. D. Genton and U. W. Kesselring, J. Pharm. Sci., 66, 676 (1977).
100. J. T. Carstensen, Drug Dev. Ind. Pharm., 14, 1927 (1988).
101. C. Ahlneck and G. Zografi, Int. J. Pharm., 62, 87 (1990).
102. T. Chen, Drug Dev. Ind. Pharm., 18, 1311 (1992).
103. M. J. Hageman, Drug. Dev. Ind. Pharm., 14, 2047 (1988).
104. *Code of Federal Regulations*, Title 21, *Food and Drugs*, Part 211, Current good manufacturing practice for finished pharmaceuticals, Subpart G, §211.137 Expiration Dating.
105. *Code of Federal Regulations*, Title 21, *Food and Drugs*, Part 211, Current good manufacturing practice for finished pharmaceuticals, Subpart I, §211.166 Stability Testing.

FRESENIUS EXHIBIT 1012
Page 160 of 293

106. *Code of Federal Regulations*, Title 21, *Food and Drugs*, Part 211, Current good manufacturing practice for finished pharmaceuticals, Subpart I, §211.170 Reserve Samples.
107. *Code of Federal Regulations*, Title 21, *Food and Drugs*, Part 211, Current good manufacturing practice for finished pharmaceuticals, Subpart J, §211.194 Laboratory Records.
108. *Code of Federal Regulations*, Title 21, *Food and Drugs*, Part 314, Applications for FDA approval to market a new drug or antibiotic drug, Subpart B, §314.50 Content and format of an application.
109. *Code of Federal Regulations*, Title 21, *Food and Drugs*, Part 314, Applications for FDA approval to market a new drug or antibiotic drug, Subpart B, §314.70 Supplements and other changes to an approved application.
110. *Code of Federal Regulations*, Title 21, *Food and Drugs*, Part 314, Applications for FDA approval to market a new drug or antibiotic drug, Subpart B, §314.81 Other postmarketing reports.
111. ICH Expert Working Group, Stability testing of new drug substances and products, International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use, October, 1993.
112. Guidelines for submitting samples and analytical data for methods validation, February, 1987, Center for Drugs and Biologics, Food and Drug Administration, Rockville, MD.
113. E. Debesis, J. P. Boehlert, T. E. Givand, and J. C. Sheridan, Pharm. Technol. 6(9), 120–137 (1982).
114. U.S. Pharmacopeial Convention, *United States Pharmacopeia, XXII*, <1225> pp. 1710–12, 1990.
115. Stability System (1989), ScienTek Software, P.O. Box 323, Tustin, CA 92681.
116. R. F. Tetzlaff, Pharm. Technol., 16(5), 70 (1992).

FRESENIUS EXHIBIT 1012
Page 161 of 293

7

Preformulation

Jens T. Carstensen
University of Wisconsin, Madison, Wisconsin

I. INTRODUCTION

Historically, preformulation evolved in the late 1950s and early 1960s as a result of a shift in emphasis in industrial pharmaceutical product development. Until the mid-1950s, the general emphasis in product development was to develop elegant dosage forms, and organoleptic considerations far outweighed such (as yet unheard of) considerations about whether a dye used in the preparation might interfere with stability or with bioavailability.

In fact, pharmacokinetics and biopharmaceutics were in their infancy and, although stability was a serious consideration, most analytical methodology was such that even gross decomposition often went undetected.

It was, in fact, improvement in analytical methods that spurred the first programs that might bear the name "preformulation." Stability-indicating methods would reveal instabilities not previously known, and reformulation of a product would be necessary. When faced with the problem of attempting to sort out the component of incompatibility in a ten-component product, one might use many labor hours. In developing new products, therefore, it would be logical to check ahead of time which incompatibilities the drug exhibited (testing it against common excipients). This way the disaster could be prevented in advance.

A further cause for the birth of preformulation was the synthetic organic programs started in many companies in the 1950s and 1960s. Pharmacological screens would show compounds to be promising, and pharmacists were faced with the task of rapid formulation. Hence, they needed a fast screen (i.e., a preformulation program) to enable them to formulate intelligently. The latter adverb implies that some of the physical chemistry had to be known, and this necessitated determination of physicochemical properties, a fact that is also part of preformulation. The approach was so logical, indeed, that it eventually became part of official requirements for INDs and NDAs [1]:

> . . . New drug substances in Phase I submission. For the drug substance, the requirement includes a description of its physical, chemical, or biological characteristics. We in the reviewing divisions regard

FRESENIUS EXHIBIT 1012
Page 162 of 293

stability as one of those characteristics. The requirement for NDA submissions . . . of the rewrite stability information is required for both the drug substance and drug product. A good time to start to accumulate information about the appropriate methodology and storage stations for use in dosage form stations for use in dosage form stability studies, therefore, is with the unformulated drug substance. . . . Stress storage conditions of light, heat and humidity are usually used for these early studies, so that the labile structures in the molecule can be quickly identified. . . . If degradation occurs, the chemical reaction kinetics of the degradation should be determined. . . . Physical changes such as changes from one polymorph to another polymorph should be examined. . . . With the drug substance stability profile thus completed, the information should be submitted in the IND submission.

II. TIMING AND GOALS OF PREFORMULATION

The goals of the program, therefore, are (a) to establish the necessary physicochemical parameters of a new drug substance, (b) to determine its kinetic rate profile, (c) to establish its physical characteristics, and (d) to establish its compatibility with common excipients.

To view these in their correct perspective, it is worthwhile to consider where (i.e., at what time) in an overall industrial program performulation takes place. The following events take place between the birth of a new drug substance and its eventual marketing (but, first of all, most investigational drug substances never make it to the marketplace for one reason or another):

1. The drug is synthesized and tested in a pharmacological screen.
2. The drug is found sufficiently interesting to warrant further study.
3. Sufficient quantity is synthesized to (a) perform initial toxicity studies, (b) do initial analytical work, and (c) do initial preformulation.
4. Once past initial toxicity, Phase I (clinical pharmacology) begins, and there is a need for actual formulations (although the dose level may not yet be determined).
5. Phase II and III clinical testing then begins, and during this phase (preferably phase II) an order of magnitude formula is finalized.
6. After completion of the foregoing, an NDA is submitted.
7. After approval of the NDA, production can start (product launch).

III. PHYSICOCHEMICAL PARAMETERS

Physicochemical studies are usually associated with great precision and accuracy and, for a new drug substance, would include (a) p*K* (if the drug substance is an acid or base), (b) solubility, (c) melting point and polymorphism, (d) vapor pressure (enthalpy of vaporization), (e) surface characteristics (surface area, particle shape, pore volume), and (f) hygroscopicity. Unlike usual physicochemical studies, an abundance of material is usually not at hand for the first preformulation studies: in fact, at the time this function starts, precious little material is supplied; therefore, the formulator will often settle for good estimates, rather than attempting to generate results with four significant figures.

There is another good reason not to aim ''too high'' in the physicochemical studies of the first sample of drug substance: Usually, the synthesis is only a first scheme; in later scale-up it will be refined and, in general, the first small samples contain some small amount of impurities, which may influence the precision of the determined constants. But it is neessary to roughly know important properties such as solubility, p*K*, and stability. These will now be dealt with in order.

FRESENIUS EXHIBIT 1012
Page 163 of 293

A. pK_a and Solubility

One important goal of the preformulation effort is to devise a method for making solutions of the drug. Frequently, the drug is not sufficiently soluble in water itself to allow the desired concentrations: for example, for injection solutions. Solubilities are determined by exposing an excess of solid to the liquid in question and assaying after equilibrium has been established. This usually is in the range of 60–72 hr, and to establish that equilibrium indeed has been attained, sampling at earlier points is necessary. Unstable solutions pose a problem here; this subject is dealt with in more detail later. Solubilities cannot be determined by precipitative methods (e.g., by solubilizing an acid in alkali and then lowering the pH to the desired pH) because of the so-called metastable (solubility) zone [2]. In the discussion to follow, drug substances are subdivided into two categories: ionizable substances, and (virtually) nonionizable substances.

Ionizable Substances

For substances that are carboxylic acids (HA), it is advantageous to determine the pK_a, since this property is of importance in a series of considerations. For carboxylic acid, the species A^- usually absorbs in the ultraviolet (UV) region, and its concentration can be determined spectrophotometrically [3]; on the other hand, HA will absorb at a different wavelength.

The molar absorbances of the two species at a given wavelength are denoted ϵ_0 and ϵ_- (and it is assumed that at the wavelength chosen $\epsilon_0 < \epsilon_-$) and it can be shown that if the solution is m_0 molar in total A, then

$$\frac{A^-}{HA} = \frac{\epsilon - \epsilon_0 m_0}{m_0 \epsilon_- - \epsilon} \tag{1}$$

so that the ratio A^-/HA can be determined in a series of buffers of different pH. Hence, the pK can be found as the intercept by plotting pH as a function of $\log[(A^-)/(HA)]$ by Henderson-Hasselbach:

$$\mathrm{pH} = \mathrm{p}K + \log(A^-/(HA)) \tag{2}$$

If several buffer concentrations are used, extrapolation can be carried out to zero ionic strength, and the pK_a can be determined. For initial studies, however, a pK in the correct range (i.e., ± 0.2 unit) will suffice, so that the foregoing determination can be performed at one buffer concentration only.

The conventional approach is to do titrations (Fig. 1), and this will yield graphs of the fraction neutralized (x) as a function of pH. Usually, the water is titrated as well [4], and what is presented in Fig. 1 is the "difference." The pK is then the pH at half neutralization (which is also the inflection point). The pH–solubility curve can now be constructed simply by determining the solubility of HA (at low pH) and A^- (e.g., of NaA) at high pH (e.g., at pH 10). Note that, at a given pH, the amount in solution in a solubility experiment is

$$S = S_{HA} + C_{A^-} \tag{3}$$

where S denotes solubility. The last term can be determined from knowledge of the pH and use of Eq. (2).

For drugs that are amines, the free base is frequently poorly soluble and, in this event, the pK is often estimated by performing the titration in a solvent containing some organic solvent (e.g., ethanol). By doing this at different organic solvent concentrations (e.g., 5, 10, 15, and 20%), extrapolation can be carried out to 0% solvent concentration to estimate the aqueous pK.

FRESENIUS EXHIBIT 1012
Page 164 of 293

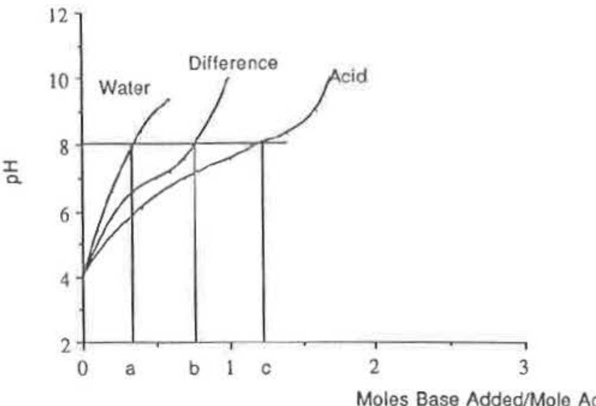


Fig. 1 Typical titration curves. The water curve indicates the amount of alkali needed to titrate the water, and the acid curve is a conventional titration curve. The difference curve is the horizontal difference between the acid and the water curve, and is the adjusted titration curve (e.g., the point b is c minus a). The pK is the point of inflection, which is also the point at which half of the acid is neutralized.

Nonionizable Substances

For hydrophobic, (virtually) nonionizable substances [i.e., those that show no ionic species of significance in the pH range 1–10 (e.g., diazepam)], solubility can usually be improved by addition of nonpolar solvents. Aside from solubility, stability is also affected by solvents, either in a favorable or in a nonfavorable direction [5]. Theoretical equations for solubility in water [6] and in binary solvents [7] have been reported in the literature, but, in general, the approach in preformulation is pseudoempirical. Most often the solubility changes as the concentration of nonpolar solvent, C_2, increases. For binary systems, it may simply be a monotonely increasing function [18], as shown in Fig. 2. The solubility is usually tied to the dielectric constant,

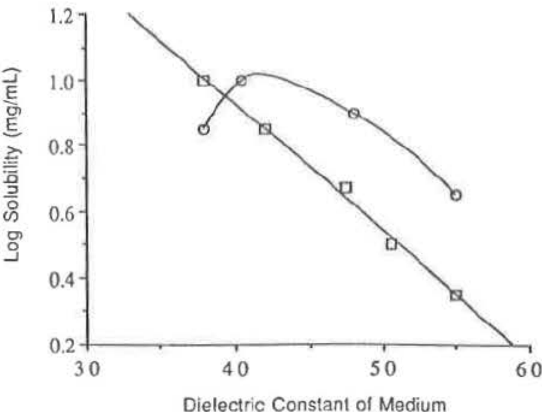


Fig. 2 (○) Solubility of 7-chloro-1,3-hydro-5-phenyl-2H-1,4-benzodiazepine-2-one-4-oxide in aqueous propylene glycol. (Data from Ref. 8.) (□) Solubility of another benzodiazepine. (Unpublished data.)

FRESENIUS EXHIBIT 1012
Page 165 of 293

and in a case, such as that shown in curve A, the solubility is often log-linear when plotted as a function of inverse dielectric constant, E, that is,

$$\ln S = \frac{-e_1}{\epsilon + e_2} \tag{4}$$

where ϵ is the dielectric constant, and the e-terms are constants [3].

Frequently, however, the solubility curve has a maximum (as shown in curve B in Fig. 2) when plotted as both a function of C_2 and ϵ [9]. In either case, it is possible to optimize solubility by selection of a solvent system with a given value of ϵ; that is, once the curve has been established, the optimum water/solvent ratio for another solvent can be calculated from known dielectric constant relationships [10].

Ternary Systems and Optimization

Frequently, *ternary* solvent systems are resorted to. Examples are water–propylene glycol–benzyl alcohol or water–propylene glycol–ethanol. In such cases the solubility profile is usually presentable by a ternary diagram [11]. This type of diagram usually demands a fair amount of work; that is, the solubility of the drug substance in many solvent compositions must be determined. A priori, it would seem, therefore, that they would be out of place in a situation where only limited quantities of drug are available. However, their principle gives some validity to optimization procedures.

The diagram can be of one of two types, as shown in Figs. 3 and 4. In the first type, the solubility may be assumed to be of the type

$$S = a_{10} + a_{11}C_1 + a_{12}C_2 \tag{5}$$

where C denotes concentrations of nonaqueous solvents. An example of this is shown in Fig. 3 inset. Here the subscripts to C denote the two nonaqueous solvents. Hence, three solubility experiments would determine the relationship (with zero degrees of freedom). It is usual to do at least five, and determine possible curvature [i.e., inclusion of more terms in Eq. (5)].

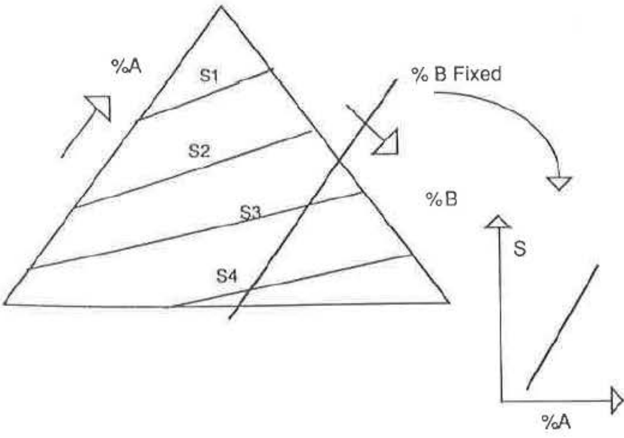


Fig. 3 Ternary diagram of solubility of a compound in a ternary mixture with linear solubility response. Inset: Concentration of drug in compositions with constant concentration of B. The composition of the solute is the constant concentration of B, the concentration of A in the abscissa, and the complement concentration of the third constituent. The drug solubility response is linear in the A-concentration here.

In the second case, in Fig. 4 component A, each tie line will give a parabolic-type curve as shown in the inset. Hence, at a given concentration of C_2 the solubility can be approximated by

$$S = b_{10} + b_{11}C_1 + b_{12}C_1^2 \tag{6}$$

where, in the simplest case,

$$b_{10} = C_{20} + C_{21}C_2 \tag{7}$$

Hence, optimization can be achieved by five (or more) experiments, with zero (or in general $n - 5$) degrees of freedom.

Prediction of Solubility

It is advantageous, with a new drug substance, to be able to estimate what its solubility might be, before carrying out dissolution experiments. There are several systems of solubility prediction, notably those published by Amidon and Yalkowsky [12–14] in the 1970s. Their equation, for solubility of *p*-aminobenzoates in polar and mixed solvents, is a simplified two-dimensional analog of the Scatchard-Hildebrand equation and is based on the product of the interfacial tension and the molecular surface area of the hydrocarbon portion of a molecule.

More recently, Bodor and Huang [15] have developed a semiempirical solubility predictor based on 20 variables (S = molecular surface in Å^2, I_a = indicator variable for alkanes, D = calculated dipole moment in Debyes; Q_n = square root of sum of squared charges on oxygen atoms; Q_o = square root of sum of squared charges on non-oxygen atoms; V = molecular volume in Å^3; S_2 = square of molecular surface; C = constant; MW = molecular weight; $\{O\}$ = ovality of molecule; A_{bh} = sum of absolute values of atomic charges on hydrogen atoms; A_{bc} = sum of absolute values of atomic charges on carbon atoms, A_m = indicator variable for aliphatic amines; and N_h = number of N-H single bonds in the molecule.

The aqueous solubilities, W, of 331 compounds followed the following equation (with tolerances omitted):

$$\begin{aligned}\log W = &-56.039 + 0.32235D - 0.59143I_a + 38.443Q_n^4 - 51.536Q_n^2 \\ &+ 18.244Q_n + 34.569Q_o^4 - 31.835Q_o^2 + 15.061Q_o + 1.9882A_m \\ &+ 0.15689N_h + 0.00014102S^2 + 0.40308S - 0.59335A_{bc} \\ &- 0.42352V + 1.3168A_{bh} + 108.80\{O\} - 61.272\{O\}^2\end{aligned} \tag{8}$$

Of the parameters listed, only the ovality and the indicator value for the alkanes I_a are unfamiliar entities that are obtained from the literature [15].

Dissolution

The importance of dissolution will be discussed in more detail later. A short note on the topic is, however, necessary for the further development at this point. According to Noyes-Whitney [16]

$$\frac{dm}{dt} = \frac{VdC}{dt} = -kA(S - C) \tag{9}$$

where

m = mass not dissolved
V = liquid volume
t = time

FRESENIUS EXHIBIT 1012
Page 167 of 293

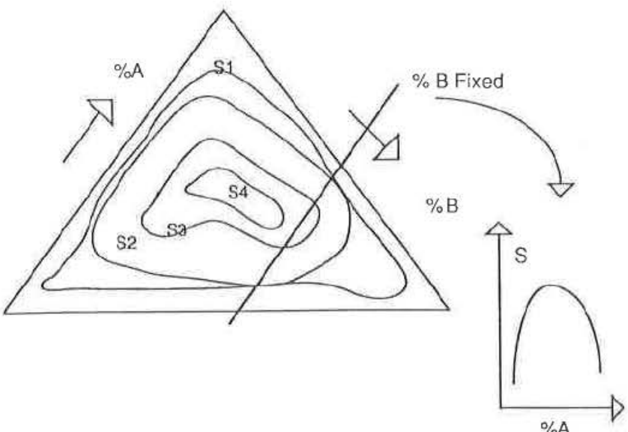


Fig. 4 Ternary diagram and tie-line concentration in a nonlinear system.

k = so-called intrinsic dissolution rate constant (cm/sec)
A = surface area of the dissolving solid

Many criticisms have been voiced against Eq. (9), but, in general, it is correct, and it will be assumed to be so in the following. If an experiment is carried out with constant surface (as e.g., using a Wood's apparatus [17], or with smaller amounts, making a small pellet and encasing it in wax, and exposing only one face to a dissolution medium, or if an excess of solid prevails throughout the dissolution experiment, then Eq. (9) may be integrated to give

$$\ln \left\{1 - \left(\frac{C}{S}\right)\right\} = \left(\frac{kA}{V}\right) t \tag{10}$$

$$C = S[1 - \exp(- \{kA/V\}\ t)] \tag{11}$$

A typical curve following Eq. (11) is shown in Fig. 5.

Solubility of Unstable Compounds

Quite often a compound is rather unstable in aqueous solution. Hence, the long exposure to liquid required for traditional solubility measurements will cause decomposition, and the resulting solubility results will be unreliable. In this particular instance, a method known as Nogami's method may be used. If a solution experiment is carried out as a dissolution experiment, with samples taken at equal time intervals, δ, it can be shown [18] that when the amount dissolved at time $t + \delta$ is plotted versus the amount dissolved at time t a straight line will ensue (Fig. 6).

Metastability is evident in dissolution rates as well. If dissolution of a metastable form is monitored at equal time intervals, δ, then, when the concentration at time $t + \delta$ is plotted versus the concentration the following relationship holds:

$$C(t + \delta) = S[1 - \exp(-k\delta)] + \exp(-k\delta)C_t \tag{12}$$

Hence, such as plot (as shown in Fig. 6) will give k from the slope, and inserting this in the intercept expression will give S. The advantage of the method is that it can be carried out in

FRESENIUS EXHIBIT 1012
Page 168 of 293

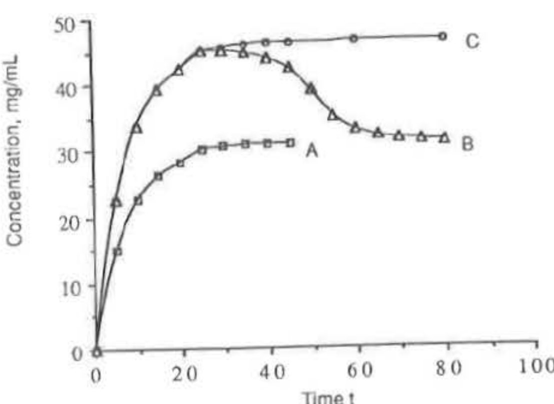


Fig. 5 Dissolution profiles obtained from the solubility determination of two polymorphic forms of the same drug substance. A is the stable form with solubility 31 mg/ml; C is the profile of the metastable form with solubility 46 mg/ml. This solubility (circles) is not achieved in many instances, and precipitation of the stable form occurs at a point beyond the solubility of A, and the trace becomes B.

a short time, and can reduce the effect of decomposition; the disadvantage is that it is not as precise as ordinary solubility determinations.

Solubility of Metastable Polymorphs

Polymorphism is an important aspect of the physical properties of drugs. One of the characteristics of a metastable polymorph (to be discussed in some detail at a later point) is that it is more soluble than its stable counterpart. The solubility profile of the drug will be as shown in Fig. 5; A is the stable form with solubility 31 mg/ml; B is the profile of the metastable form with solubility 46 mg/ml. This solubility (circle) is usually not achieved, and precipitation of the stable form occurs at a point beyond the solubility of A, and the trace becomes B.

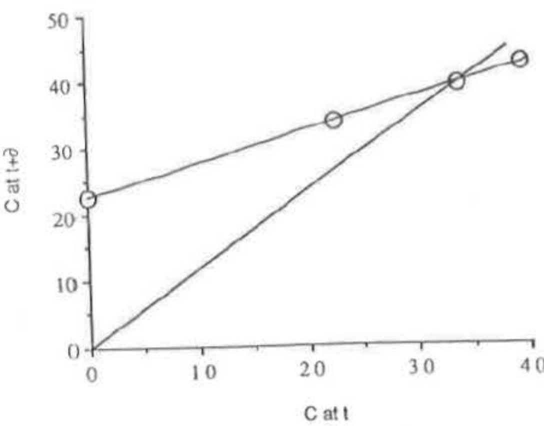


Fig. 6 The Nogami method applied to the data in Fig. 5.

FRESENIUS EXHIBIT 1012
Page 169 of 293

In such cases, the Nogami method can be applied to the early points curve (see Fig. 6), and the solubility S', of the polymorph can be assessed. One of the important aspects of metastable polymorphs in pharmacy is exactly their higher solubility, since the dissolution rate will also be higher [see Eq. (9)]. Hence, the bioavailability will be increased where this is dissolution rate-limited [19].

Polymorphism

Solids exist as either amorphous compounds or as crystalline compounds. In the latter, the molecules are positioned in lattice sites. A lattice is a three-dimensional array, and there are eight systems known. Compounds often have the capability of existing in more than one crystalline form, and this phenomenon is referred to as *polymorphism*.

If a compound exhibits polymorphism, one of the forms will be more stable (physically) than the other forms; that is, of n existing forms, $n - 1$ forms will possess thermodynamic tendency to convert to the nth, stable form (which then has the lowest Gibbs energy; however, in the preformulations stage, it is not known whether or not the form on hand is the stable polymorph).

One manner in which different polymorphs are created is by way of recrystallizing them from different solvents, and at a time point when sufficiently material (and this need not be very much) is available, the preformulation scientist should undertake recrystallization from a series of solvents.

Knowledge of polymorphic forms is of importance in preformulation because suspension systems should never be made with a metastable form (i.e., a form other than the stable crystalline form). Conversely, a metastable form is more soluble than a stable modification, and this can be of advantage in dissolution [see Eq. (11)]. There are two types of polymorphism, a fact illustrated in the following discussion.

If the vapor pressure or solubility of a compound is plotted as a function of temperature, a plot, such as that shown in Fig. 7, will result. Here form I is the form that is stable at temperature T_1. If the compound exists in the two forms I and II, the phenomenon is referred to as an enantiotropic system, since, on heating to the temperature T_2, form I will transform into form II. Form II may exist below temperature T_2, but perturbations (e.g., presence of moisture) will convert it to form I, and the energy involved in the transformation will be

$$E = RT \ln\left[\frac{S_2}{S_1}\right] \tag{13}$$

If the compound is present as form I at room temperature and heated up fast it will melt at T', which is lower than the melting point of form II (T'').

A different situation exists if the compound exists as form I and form III. This is referred to as a monotropic system, and here, III is unstable relative to I over the whole solid range. In this case, however, the melting point of the "unstable" polymorph is lower than that of the stable one (T''' is lower than T').

If a metastable polymorph is kept dry, it may be stable for eons and, therefore, it is not referred to as an "unstable," but rather, as "metastable". An even more energetic state is represented by amorphous forms, which may be considered supercooled liquids. Today, polymorphism is checked for in two fashions. Thermal methods will give information about whether a polymorph is stable, enantiotropic, or monotropic. If the system possesses a transition point, G is zero at this point. In a fashion similar to the melting process, where G is also zero, the transition is associated with an enthalpy change, which is endothermic. Hence, if a sample of a compound is heated up and a change in enthalpy occurs below the melting point, it is an enantiotropic transformation if the point is reproducible and if it occurs again after the sample

FRESENIUS EXHIBIT 1012
Page 170 of 293

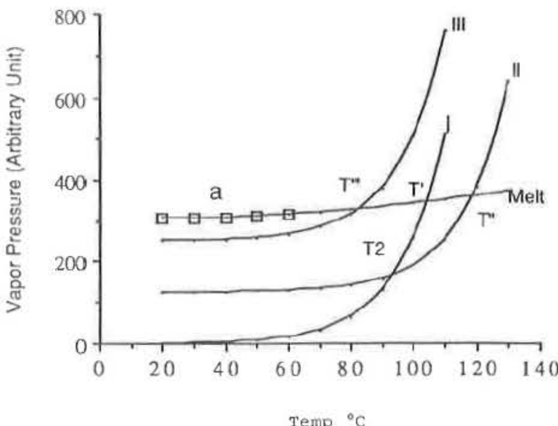


Fig. 7 Dependence of vapor pressure and solubility for an enantiotropic pair (I/II), a monotropic pair (I/III), and an amorphous compound (a).

has been cooled back down and shows the same enthalpy change on heating. Otherwise, it is monotropic and, in this case, as shown in Fig. 8, there will either be a (lower) melting point (T''') with a single endotherm, or there will be a melting followed by a recrystallization into the more stable forms that then melt [denoted III (Alternate)] in Fig. 8.

An amorphous compound has no melting point, and (above glass transition temperatures) its vapor pressure curve simply continues into that of the melt.

Second, x-ray diffraction will give spacings directly, in the crystal and reveal differences between samples. Finally, solubility curves can be carried out and, if a nick in the solubility curve is found (Fig. 9), this is a transition temperature. If no nick is found, there is no transition temperature, but if the dissolution curves are as shown in Fig. 5, there is polymorphism and, by indirect argumentation, it may be concluded that a monotropic system exists.

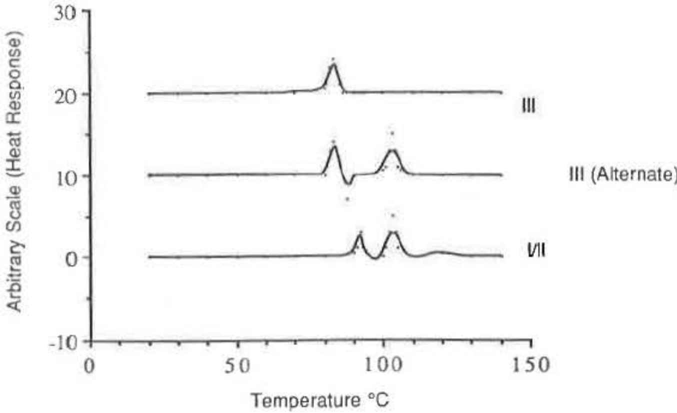


Fig. 8 DSC tracings of the polymorphs shown in Fig. 7.

FRESENIUS EXHIBIT 1012
Page 171 of 293

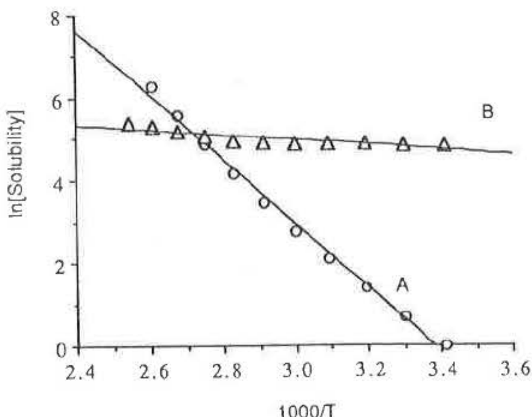


Fig. 9 Solubility curves of an enantiotropic pair.

C. Partition Coefficient

Partition coefficients between water and an alkanol (e.g., octanol) should be determined in preformulation programs [20]. The partition coefficient of a compound that exists as a monomer in two solvents is given by

$$K = \frac{C_1}{C_2} \tag{14}$$

If it exists as an n-mer in one of the phases, the equation becomes

$$K = \frac{(C_1)^n}{C_2} \tag{15}$$

or

$$\log k = n \log C_1 - \log C_2 \tag{16}$$

The easiest way to determine the partition coefficient is to extract V_1 ml of saturated aqueous solution with V_2 ml of solvent, and determine the concentration C_2 in the latter. The amount left in the aqueous phase is $(C_1V_1 - C_2V_2) = M$, so that the partition coefficient is given by

$$K = \frac{M}{V_1C_2} \tag{17}$$

If it is assumed that the species is monomeric in both phases, the partition coefficient becomes the ratio of the solubilities, and it is simply sufficient to determine the solubility of the drug substance in the solvent (since it is assumed that the solubility is already known in water):

$$K = \frac{S_1}{S_2} \tag{18}$$

D. Vapor Pressure

In general, vapor pressures are not all that important in preformulation, but it should always be kept in mind that a substance may have a sufficiently low vapor pressure that it will (a)

FRESENIUS EXHIBIT 1012
Page 172 of 293

become lost to a large enough extent that apparent stability and content uniformity problems will result; and (b) it will exhibit a potential for interaction with other compounds and adsorption onto or sorption into package components [21].

Most drug substances are not substantially volatile. As an initial screen, it can be determined whether the drug is sufficiently volatile to cause concern by placing a weighed amount of it in a vacuum desiccator and weighing it daily for a time. It is better to have a high-vacuum system for this, and the use of a vacuum electrobalance is best for this purpose. A good estimate of the vapor pressure can be obtained [22] by using a pierced thermal analysis cell, placing it on a vacuum electrobalance, and monitoring the rate of weight loss. A substance with known vapor pressure can then be used for calibration, the loss rates being proportional to the vapor pressures.

E. Surface Characteristics

The surface characteristics of a batch of a drug substance may greatly influence its properties in processing (flow, dissolution). Crystals may form in different habits (plate, needle, cube) and these may not be due to morphology; that is, depending on crystallization circumstances, they could all be the same crystal form, but of different habit [23].

It is a good practice, both during development of a new drug and through to the NDA, to take photomicrographs of each new batch of drug substance delivered to the product development department. In this manner, there will be a permanent reference record, and when deviations from expected behavior occur during the product development sequence, the photomicrograph will be one record that may throw light on the problem. Aside from this, the specific surface area (A'', cm^2/g) of each batch of drug substance should be measured.

Shape and Fractal Dimension

Shape is of great interest and affects many properties; therefore, it is important to have a record of how a shape changes as the synthesis of the raw material undergoes changes during the developmental process. In the simplest form, microscopy of all batches used in product development should be carried out to determine the ratio of longest to shortest dimension (average of ten measurements). This is a type of shape factor.

However, there are more sophisticated methods that may be used to attain a good feel for the shape factor; namely, its fractal dimension. This is most conveniently carried out by use of imaging techniques.* The general principle of this is shown in Figs. 10 and 11.

As an example (and this is a hypothetical example only), a particle is shown in Fig. 10, such as it might appear on a microscope slide. This particle is gridded out in the form shown in the grid in the upper left-hand corner of Fig. 10. The number of squares in which parts of the trace of the particle are located is counted. This number is N, and the length of the grid size is g. The grid size is arbitrarily set equal to 1 in this example. The grid length is now halved, and the number of squares counted again; this is then repeated several times. The fractal equation is then given by

$$\ln[N] = -n \ln[g] + q \tag{19}$$

*An example of an excellent system is that of Universal Imaging Concepts Image 1 Software 486/66 computer (Data Store, Inc); Hamamatsu C2400-77H B&W chip camera (CCD Camera Control); Sony 19" PVM1943MD color monitor; UP5000 (Sony) color printer; Nikon model Labophot-2 40x ocular x 10 on-screen magnification; Nikon polarizing microscope, model POH-II.

FRESENIUS EXHIBIT 1012
Page 173 of 293

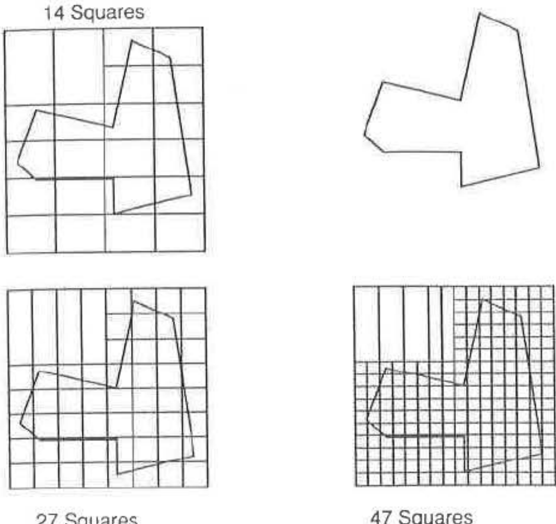


Fig. 10 Dividing the field containing a particle into more and more squares, by dividing the original grid length, a, by 2.0 in each step. The length would be $L = Na\sqrt{2}/g$ if it were a line going through corners of the squares, so that if it were a straight line, then $\ln[L]$ should equal $\ln[N] - \ln[g] + \ln[a\sqrt{2}]$, but the slope differs from one that is the "natural" dimension of a line.

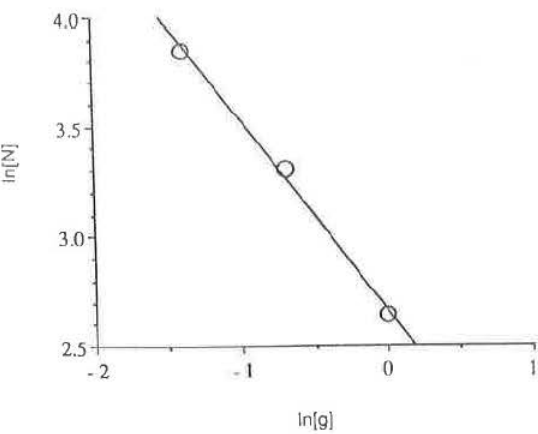


Fig. 11 Fractal graph of the data in Fig. 10. The least squares fit for the equation is $\ln[N] = 2.6575 - 0.8735 \ln[g]$.

FRESENIUS EXHIBIT 1012
Page 174 of 293

where n and q are constants, and where $n + 1$, the slope of the line is the fractal dimension [24]. This dimension is characteristic for shapes, and its determination is worthwhile to keep record of how shape may change with synthesis changes.

The number of decreases in the so-called measuring stick (here, the grid) is of great importance, and it is the behavior at the smaller end of the measuring stick size that is important. In fact the small number of iterations leading to Eq. (19) may well be fallacious but at the lower end of the scale the relationship will still be correct (but the true fractal dimension may or will be different from the value found in Fig. 10).

Particle Size and Size Distributions

The "size" of the particles is of great importance, and particle size determinations should be carried out in preformulation as well in formulation functions. For small particle sizes, simple microscopy [25] may be used, but again, imaging techniques, particularly with motorized stages, are more representative and much easier to carry out [26].

When characterizing a powder, it is important to make certain what type of diameter is being described, since various techniques give different "types of diameter."

The volume-surface mean diameter, d_{sv}, can be determined by permeamitry (e.g., Fisher Sub-Sieve Sizer) for a fine powder [25]. It is given by

$$d_{sv} = \frac{6\ V}{A''} \tag{20}$$

where V is the solids volume of the sample and A'' is the geometric surface area. A'' can be obtained for a coarse powder by sieve analysis:

$$A'' = \{6/\rho\}\ \Sigma w_i/d_i \tag{21}$$

where ρ is the solids density (g/cm^3), d_i the mean diameter of the ith mesh cut, and w_i the weight retained by the ith screen. If the drug substance is not porous, the ratio

$$G = \frac{A}{A''} \tag{22}$$

is the rugosity and is a measure of the surface roughness.

Pore Size Distribution

With the advent of mercury intrusion porosimeters, it is advantageous to perform a pore size distribution of investigational batches of a drug [27]. The Washburn equation [28] states that the pressure P necessary to intrude a pore is given by

$$\ln P = -q\gamma\left(\frac{1}{r}\right)\cos\,[\theta] \tag{23}$$

where

q = a constant
γ = interfacial tension
θ = contact angle
r = radius of the pore being penetrated

Since mercury has a contact angle with most solids of about 140°, it follows that its cosine is negative (i.e., it takes applied pressure to introduce mercury into a pore). In a mercury porosimeter, a solid sample is evacuated in a cell, mercury is then intruded, the volume V is noted (it actually reads out), and the pressure P is then increased stepwise. In this fashion it

is possible to deduce the pore volume of a particular radius [corresponding to P by Eq. (23)]. A pore size distribution will give the total internal pore area as well, which can be of importance in dissolution.

Hygroscopicity

Hygroscopicity is an important characteristic of a powder. It can be shown, roughly, for a fairly soluble compound that the hygroscopicity is related to its solubility [29,30], although the heat of solution plays an important part in what is conceived as "hygroscopicity" [31–33]. A hygroscopicity experiment is carried out most easily by exposing the drug substance to an atmosphere of a known relative humidity (e.g., storing it over saturated salt solutions in desiccators). Each solution will give a certain relative humidity (RH), and the test is simply to weigh the powder from time to time and determine the amount of moisture absorbed (weight gained). This does not work with drug substances that decompose (e.g., effervescent mixtures will start losing weight owing to carbon dioxide evolution [34]).

If the air space is agitated sufficiently to prevent vapor pressure gradients, the initial uptake rate (g, H_2O/g solid per hour) is related to relative humidity by

$$L = a_{21}[RH - RH_0] \tag{24}$$

where RH_0 is the vapor pressure of a saturated solution of the drug substance in water. The latter can be estimated by an ideality assumption; that is, if the solubility is expressed as a mole fraction X_s, the vapor pressure over a saturated solution will be P' given by

$$P' = (1 - X_s)P^* \tag{25}$$

where P^* is water's vapor pressure at that temperature.

The foregoing experiments are rather easy to carry out and should always be part of a preformulation program, since hygroscopicity can be so important that it will dictate whether or not a particular salt should be used. Flurazepam (Dalmane), for instance, is a monosulfate, and is used as such, since the disulfate, desirable in many other respects, is so hygroscopic that it will remove water from a hard-shell capsule and make it exceedingly brittle.

IV. COMPATIBILITY TESTS

It should again be emphasized that, at the onset of a new drug program, there are only small amounts of drug substance at hand. One of the first tasks for the preformulation scientist is to establish the framework within which the first clinical batches can be formulated. To this end, it is important to know with which common excipients the drug is compatible. The distinction will be made between solid and liquid dosage forms in the following sections.

A. Compatibility Test for Solid Dosage Forms

It is customary to make a small mix of drug substance with an excipient, place it in a vial, place a rubber stopper in the vial, and dip the stopper in molten carnauba wax (to render it hermetically sealed). The wax will harden and form a moisture barrier up to 70°C. A list of common excipients characteristic of this type of test is shown in Table 1. At times, it is possible to obtain quantitative relationships to excipient characteristics and interaction rates [35,36].

In addition to the test, as described, a similar set of samples are set up to which 5% moisture is added. A storage period of 2 weeks at 55°C (except for stearic acid and dicalcium phosphate, for which 40°C is used) is employed, after which time the sample is observed physically for (a) caking, (b) liquefaction, (c) discoloration, and (d) odor or gas formation. It is then assayed by thin-layer chromatography (or HPLC).

Table 1 Categories for Two-Component Systems

Additive		Identical	17–27 mo at 25°C	Worse 10 days at 55°C	Total 25°C	Score 55°C
Drug per se	Dry	15	4	1	38	31
	5% H_2O	9	8	3	49	38
+ Magnesium stearate	Dry	16	3	1	34	30
	5% H_2O	15	4	1	43	35
+ Calcium stearate	Dry	13	4	3	37	32
	5% H_2O	12	5	3	38	35
+ Stearic acid	Dry	15	5	0	42	31
	5% H_2O	7	11	2	60	38
+ Talc	Dry	14	5	1	38	30
	5% H_2O	10	8	2	45	34
+ Acid-washed talc	Dry	12	8	0	44	31
	5% H_2O	10	9	1	49	35
+ Lactose	Dry	12	5	3	38	32
	5% H_2O	9	7	4	65	56
+ $CaHPO_4$, anhydrous	Dry	12	6	2	46	36
	5% H_2O	9	8	3	66	53
+ Cornstarch	Dry	12	5	3	39	34
	5% H_2O	10	5	5	40	37
+ Mannitol	Dry	10	7	3	39	31
	5% H_2O	8	7	5	47	45
+ Terra alba	Dry	14	6	0	41	28
	5% H_2O	11	6	3	50	45
+ Sugar 4x	Dry	12	6	2	41	34
	5% H_2O	9	7	4	63	61

Source: Ref. 35.

Note that one of the samples set up is the drug by itself. This is done for several reasons, one of which is that it is now required by the FDA for IND submissions [1]. One more reason is that, at the onset of a program, the organic synthesis of the compound may lack the refinement it will later have, and it is not uncommon that there will be several weak spots (impurities) on a TLC chromatogram of a compound obtained by initial laboratory synthesis. Hence, in selecting the excipients with which the drug substance is deemed to be compatible, it is customary to use as criteria that (after accelerated exposure of a drug–excipient mix) no new spots have developed, and that the intensity of the spots in the drug that has been stored under similar conditions (2 weeks at 55°C) are the same as in the acceptable excipient. This type of program is used by many companies with good success (i.e., the formulas developed based on the findings from the compatibility program are stable).

Liquefaction occurs at times because of eutectic formation (e.g., often with caffeine combinations), and this may not necessarily be associated with decomposition. On the other hand, discoloration (e.g., amines and sugars) usually is.

Finally, the reason for not forcing dicalcium phosphate (a very valuable formulation aid in direct compression) beyond 40°C is that at higher temperatures (actually above 70°C [37]) it converts to the anhydrate, a conversion that is, curiously enough, catalyzed by water. In other words, the dihydrate will be autocatalytic at elevated temperatures, and it should not be ruled out based on high-temperature findings.

B. Kinetic pH Profiles

Frequently, a broad screen of stability is performed on the first small sample used for initial preformulation; this is frequently referred to as "forced decomposition studies" [38]. In this the drug is exposed to "acid degradation," "base degradation," "aqueous degradation," "drug powder degradation," and "light degradation." More refined studies are eventually needed.

For any compound marketed by a pharmaceutical concern, at one time during its development, there should be a concerted project to establish a very exact pH profile. To do this correctly is a time-consuming undertaking. However, the information that can be gleaned from it is very important for formulations; therefore, it is customary to carry out an approximate kinetic pH profile [39] early in the developmental stage. This will allow formulation of solutions for injections, and for oral products as well, at a pH and with buffers that will give the best stability. Without it, formulation is essentially guesswork.

Most drug decompositions are hydrolyses, during which the drug concentration C decreases with time through

$$\frac{dC}{dt} = -k_2 C_{H_2O} C \tag{26}$$

Since the water concentration hardly changes, this (bimolecular) reaction scheme is reduced to the pseudo-first-order expression

$$\frac{dC}{dt} = -KC \tag{27}$$

where K is the first-order rate constant. This integrates to the well-known form

$$\ln\left[\frac{C}{C_0}\right] = -Kt \tag{28}$$

Hence, semilog plotting of concentration versus time (see Fig. 12) will give a straight line, with a slope from which K is calculated. But, most reactions are catalyzed by buffers, by hydrogen ions, and by hydroxyl ions, so K will be of the form

$$K = k + k_+(H^+) + k_-(OH^-) + k_B(B) \tag{29}$$

where B denotes buffer concentration and k is a rate constant. A decomposition experiment is now carried out at, say, five pH values, each using two buffer concentrations. A graph is drawn of k (Fig. 13) versus B at the various pH values. A line is drawn through the two at low pH (at which hydroxyl ion concentration can be disregarded), and Eq. (29) becomes

$$K = k_+(H^+) + k_B(B) \tag{30}$$

FRESENIUS EXHIBIT 1012
Page 178 of 293

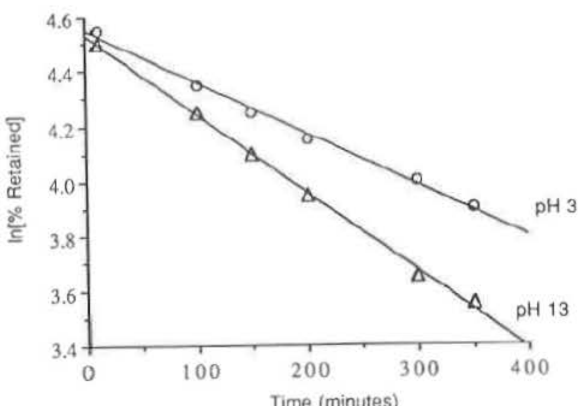


Fig. 12 Pseudo-first-order decompositions of carbuterol at 85°C at an ionic strength of 0.5. (Graphs plotted from data in Ref. 40.)

so that the plot in Fig. 13 will give

$$K = k_+(\mathrm{H}^+) \tag{31}$$

as intercept and k_B as slope. The foregoing allows assessment of the effect of the buffer (which is an important point in the buffer selection). Taking the 10 log of Eq. (31) now gives

$$\log K = -\mathrm{pH} + \log(k_+) \tag{32}$$

A similar argument will show that at high pH

$$\log K = \mathrm{pH} - 14 + \log(k_-) \tag{33}$$

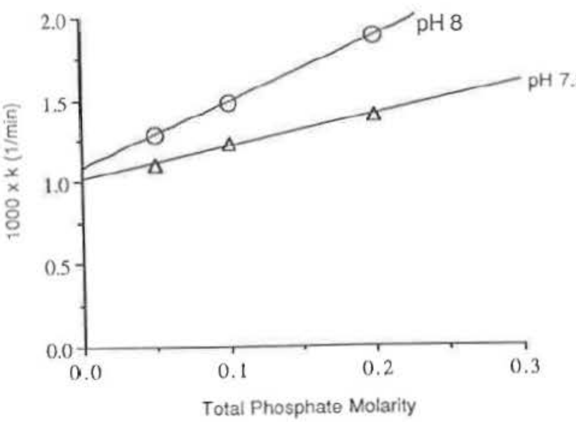


Fig. 13 Buffer concentration dependence of carbuterol at 85°C at an ionic strength of 0.5. (Graphs plotted from data in Ref. 40.)

This explains why the extremes of a pH profile (see Fig. 14) often have slopes of plus or minus unity [40]. The horizontal part is due to the uncatalyzed rate constant, k_{+}, in Eq. (29). A pH profile can be done at, for example, six pH values and, since there are two kinetic points (times) and two buffer concentrations at each, a total of 24 assays are needed, which is not insurmountable. This number may be minimized and optimized by careful selection of pH and buffer concentrations [39]. Later in the program the pH profile should be repeated, but with multiple points and several buffer concentrations, but this is beyond the point of preformulation.

C. Liquid Compatibilities

The pH profile is the most important part of liquid compatibilities. However, two-component systems are set up in aqueous (or other types of) solutions and treated as in Sec. IV.A. This is required in the 1987 stability guidelines [1], which state that "it is suggested that the following conditions . . . be evaluated in studies on solutions or suspensions of bulk drug substances: acidic and alkaline pH, high oxygen and nitrogen atmospheres, and the presence of added substances, such as chelating agents and stabilizers," and it is suggested "that stress testing conditions . . . include variable temperature (e.g., 5, 50, and 75°C)."

Aqueous Solution Capability

In general, such studies are carried out by placing the drug in a solution of the additive. This can be (and usually is) a heavy metal (with or without chelating agents present) or an antioxidant (in either oxygen or nitrogen atmosphere). Usually, both flint and amber vials are used and, in many cases, an autoclaved condition is included. This will answer questions about susceptibility to oxidation, to light exposure, and to heavy metals. These are important questions as far as injectable compatibilities are concerned. Exposure to various plugs is frequently included at this point so that early injectable preparations can be formulated.

For preparations for oral use, knowledge of the desired dosage form is important, but compatibility with ethanol, glycerin, sucrose, corn syrup, preservatives, and buffers is usually carried out. This type of study also gives an idea of the activation energy E of the predominant reaction in solution. The Arrhenius plots (Fig. 15) for compounds in solution are usually quite

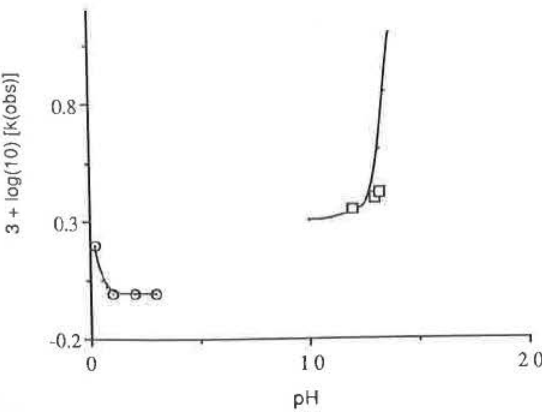


Fig. 14 Kinetic pH profile at low and high pH of carbuterol at 85°C. (Graphs plotted from data in Ref. 40.)

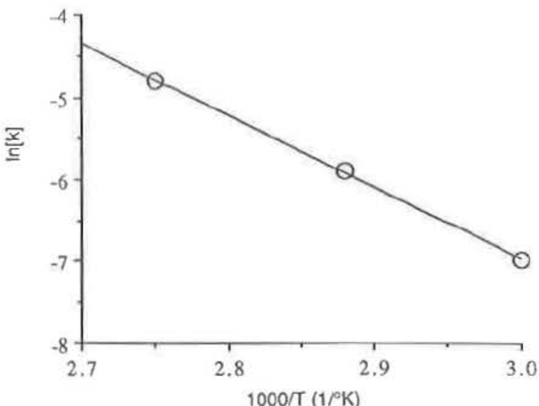


Fig. 15 Arrhenius plotting of carbuterol rate constants at pH 4 of carbuterol. (Graph plotted from data in Ref. 40.)

precise; that is, denoting the rate constant by k, absolute temperature by T, the gas constant by R, and a collision factor by Z:

$$k = Z^{-E/(RT)} \tag{34}$$

or its more useful logarithmic cousin,

$$\ln k = \ln Z - \left(\frac{E}{R}\right)\left(\frac{1}{T}\right) \tag{35}$$

Nonaqueous Liquid

With transdermal dosage forms being of great importance of late, it is advisable to test for compatibilities with "ointment" excipients and with polymers (e.g., ethylvinyl polymer, if that is the desired barrier). With transdermals, the dosage form is either directly placed in a stirred liquid, or it is placed in a cell with an appropriate membrane (e.g., cadaver skin) to estimate the release characteristics of the drug from the ointment [41].

If the overall flux is J, then

$$\frac{1}{J} = \left[\frac{1}{J_{\text{ointment}}}\right] + \left[\frac{1}{J_{\text{membrane}}}\right] \tag{36}$$

where subscripts refer to the respective phase. J_{membrane} can be obtained from curves, such as that shown in Fig. 16, in the fashion that the overall flux is first obtained (with the membrane in place), giving the value of J. Then the release is obtained without the membrane in place, giving J_{ointment}, that is,

$$J = \left(\frac{1}{A}\right)\left(\frac{dm_1}{dt}\right) \tag{37}$$

and

$$J_{\text{ointment}} = \left(\frac{1}{A}\right)\left(\frac{dm_2}{dt}\right) \tag{38}$$

J_{membrane} is then obtained as the reciprocal of the difference.

FRESENIUS EXHIBIT 1012
Page 181 of 293

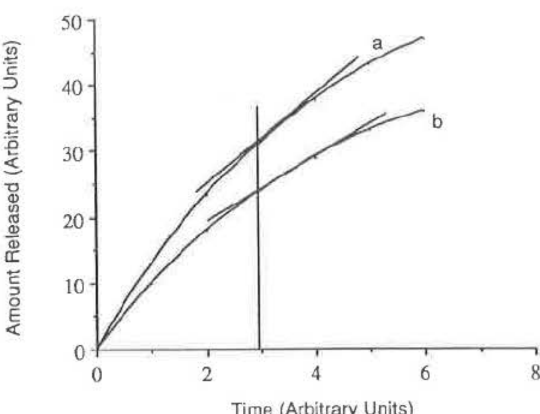


Fig. 16 Slope determination of flux of ointment release and release from ointment + membrane.

In vivo testing is usually carried out by applying the dosage form to hairless rats followed by subsequent sacrifice. Since the skin consists of several layers with differing hydrophilicity, the overall fate of the drug is of importance.

V. DISSOLUTION OF DRUG SUBSTANCE AND DOSAGE FORM

In the time path, solid dosage forms (tablets or capsules) must eventually be manufactured for the clinic (e.g., in Phase II). If possible, the drug substance per se is subjected to a dissolution test in a Wood's apparatus [17]. This test is useful, although quite dependent on hydrodynamic conditions. It consists of placing the powder in a type of tablet die, compressing the tablet, and exposing the flat, exposed side of the tablet (with surface area A) to a dissolution liquid (usually water or N/10 HCl) in which it has a solubility S. Under these conditions [42] the intrinsic dissolution rate constant (cm/sec) can be obtained by Eq. (9), which under sink conditions* (i.e., where C is less than 15% of S) becomes

$$C = \left(\frac{SkA}{V}\right)t \tag{39}$$

It has been suggested [43] that if k is obtained under sink conditions over a pH range of 1–8 at 37°C in a USP vessel by way of Eq. (39) at 50 rpm, then if the dissolution rate constant (kA/V) is greater that 1 mg min^{-1} cm^{-2}, the drug is not prone to give dissolution–rate-limited absorption problems. On the other hand, if the value is less than 0.1, such problems can definitely be anticipated, and compounds with values of kA/V of from 0.1 to 1 mg min^{-1} cm^{-2} are in a gray area. For compound selectivity it is frequently useful to express dissolution findings in terms of k (i.e., in cm/sec).

For a small amount of powder, dissolution of the particulate material can often be assessed (and compared with that of other compounds) by placing the powder in a calorimeter [44] and

*Strictly speaking, sink conditions are when the amount dissolved plotted versus time yields a line that, within experimental error, is linear. When the surface area, A, is constant, then this corresponds to 15% dissolved. When the surface area changes (e.g., during particulate dissolution), then this number may be smaller.

FRESENIUS EXHIBIT 1012
Page 182 of 293

measuring the heat evolved as a function of time. The surface area must be assessed microscopically (or by image analyzer), and the data must be plotted by a cube root equation [45]:

$$1 - \left[\frac{M}{M_0}\right]^{1/3} = \left(\frac{2kS}{\rho r}\right)t \tag{40}$$

where M is mass not dissolved; M_0, the initial amount subjected to dissolution, ρ is true density; S is solubility, and r is the mean "radius" of the particle. The method is simply comparative, not absolute, owing to the hydrodynamics being different in the calorimeter than it would be in a dissolution apparatus.

A. Biopharmaceutical Aspects

One important aspect of drug dosage form development is to obtain a dosage form that is absorbed in a desired fashion. This usually implies a rapidly and completely absorbed dosage form. This means that it is necessary to test the drug substance itself for in-vivo release characteristics. A good indication of whether a drug may give this type of problems is a comparison of LD_{50} values by the parenteral and by the oral route. If the former toxicity is much greater than the latter, there is often an absorption problem. In the following, it is assumed that the problems are dissolution dictated.

B. Partial-In Vivo Testing

The general goal is to submit an IND for the drug and to get it into testing in humans in the clinic. Frequently, biological absorption characteristics are checked by such procedures as the everted sac technique [46]. Here, a segment of the small intestine of a rat is everted. The ends are then tied off, and physiological fluid containing no drug (placebo liquid) is filled into the sac. It is then placed in a vessel containing a solution of the drug in a buffer solution. The setup is kept at 37°C, while oxygen is constantly being supplied to the solution. After a given interval, the contents of the sac are assayed for content of drug. This can then be repeated for other times. Collections of several samples from the same intestine segment is possible [47], and this greatly facilitates the procedure. Other methods, such as the method suggested by Dolusio et al. [48], exist and are used in preformulation efforts. Here, rats are anesthetized, and the ileal and duodenal ends of their intestines are then cannulized, allowing sampling and liquid introduction.

C. In Vivo Testing

For preformulation purposes, some animal testing is usually performed before Phase I. This could be in rats, dogs, or other species. The animals are being tested by being given a specific regimen (e.g., fasting, single-dose, or after each meal), and blood is then collected at various intervals. In this fashion a blood-level curve is obtained (Fig. 17). The ultimate value of this is not an absolute. To extrapolate from one species to another is a dubious undertaking, and the ramifications of this are beyond the scope of the chapter. General conclusions can be drawn, and methods are briefly described in the following.

It is customary to do one set of tests by the parenteral route and, at this stage (see Fig. 17), to assume or define that 100% of the dose is absorbed [49]. Depending on pathways, this may later be modified, but the correctness of the assumption is less important now, as will be seen shortly. Next, a solution (or if solubility is limited, a suspension) is administered, giving a different blood level (see curve II in Fig. 17). Finally, the drug is administered dry (e.g., in a capsule), giving rise to curve III in Fig. 17.

FRESENIUS EXHIBIT 1012
Page 183 of 293

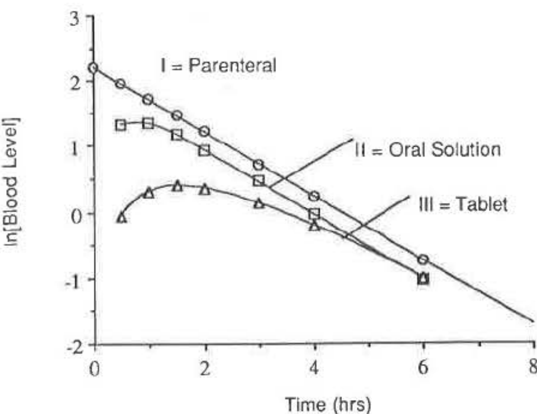


Fig. 17 Blood level curves of identical doses of a drug by parenteral route (I), by oral route as a solution (II), and orally as a tablet (III).

If the areas under the non-logarithmic curves are denoted by A, then $A\mathrm{II}/A\mathrm{I}$ is the fraction absorbed by oral route. $A\mathrm{III}/A\mathrm{II}$ is the fraction efficiency of the solid dosage form. The latter is because the solid dosage form has to dissolve before the drug contained in it is available for absorption. It is the latter ratio that is important to the investigating pharmaceutist; therefore, the outcome of the parenteral form is actually not a consideration from a formulations point of view. It is critical overall and, if it is low, it may, at the point of parenteral data acquisition, be advisable to stop the program and evaluate the possibility of derivatives that would give better availability.

There are large volumes of literature written on this subject, but the simplest manner of evaluating the formula efficiency by way of blood level data is the following simplified model: It is assumed that on ingestion the tablets or capsules (a) disintegrate into (N) particles, and (b) the drug then dissolves from these particles [50,51]. Each particle will release its content in T time units. The number of particles not dislodged from the dosage form at time t is

$$N_t = N_0 e^{-kt} \tag{41}$$

For this type of situation the blood levels B will be a function of time by an equation of the type

$$B = A[\exp(-k_1 t) - \exp(-k_2 t)] \tag{42}$$

where k_1 is an elimination rate constant, and k_2 is a function of dissolution and absorption rate constants. Figure 17 shows a case in point, where

$$B\mathrm{II} = 7[\exp(-0-494t) - \exp(-2-813t)] \tag{43}$$

$$B\mathrm{III} = 4.48[\exp(-0.472t) - \exp(-0.85t)] \tag{44}$$

The k_1 values are about the same as they should be, and the difference, k'', between the k_2 values,

$$k'' = 2.81 - 0.85 = 2.0 \tag{45}$$

FRESENIUS EXHIBIT 1012
Page 184 of 293

is a parameter indicative of the retardation-caused disintegration and dissolution. The areas under the non-logarithmic curves, Q, are

$$Q = \left(\frac{1}{k_1}\right) - \left(\frac{1}{k_2}\right) \tag{46}$$

That is, in Fig. 17, the areas would be 0.942 μg hr/cm^3 for the tablet and 1.67 for the solution. The formulation efficiency is 0.94/1.67 = 56% by this criterion, and the tablet formulation could stand some improvement.

REFERENCES

1. R. C. Schultz, Stability of dosage forms, FDA–Industry Interface Meeting, Washington, DC, Oct. 7, 1983; Stability guidelines, *Congressional Record*, May 7, 1984.
2. N. Rodriguez-Hornedo, Ph.D. Thesis, University of Wisconsin (1984).
3. W. J. W. Underberg and H. Lingeman, J. Pharm. Sci., 72, 553 (1983).
4. T. Parke and W. Davis, Anal. Chem., 25, 642 (1954).
5. S. K. Bakar and S. Niazi, J. Pharm. Sci., 72, 1024 (1983).
6. S. H. Yalkowsky and S. C. Valvani, J. Pharm. Sci., 72, 912 (1983).
7. W. E. Acree and J. H. Rytting, J. Pharm. Sci., 72, 293 (1983).
8. J. T. Carstensen, K. S. Su, P. Maddrell, and H. Newmark, Bull. Parenter. Drug Assoc., 25, 193 (1971).
9. A. N. Paruta and S. A. Irani, J. Pharm. Sci., 54, 1334 (1964).
10. G. Cave, F. Puisieux, and J. T. Carstensen, J. Pharm. Sci., 68, 424 (1979).
11. D. Scroby, R. Bitter, and J. Webb, J. Pharm. Sci., 52, 1149 (1963).
12. S. H. Yalkowsky, G. L. Flynn, and G. L. Amidon, J. Pharm. Sci., 61, 983 (1972).
13. G. L. Amidon, S. H. Yalkoskky, and S. Leung, J. Pharm. Sci., 63, 1858 (1974).
14. S. H. Yalkowsky, G. L. Amidon, G. Zografi, and G. L. Flynn, J. Pharm. Sci., 64, 48 (1975).
15. N. Bodor, Z. Gabanyi, and C.-K. Wong, J. Am. Chem. Soc., 111, 3783 (1989).
16. A. Noyes and W. Whitney, J. Am. Chem. Soc., 23, 689 (1897).
17. J. H. Wood, G. Catacalos, and S. Lieberman, J. Pharm. Sci., 52, 296 (1963).
18. H. Nogami, T. Nagai, and A. Suzuki, Chem. Pharm. Bull., 14, 329 (1966).
19. M. Shibata, H. Kokobu, K. Morimoto, K. Morisaka, T. Ishida, and M. Inoue, J. Pharm. Sci., 72, 1436 (1983).
20. S. H. Yalkowsky, S. C. Valvani, and T. J. Roseman, J. Pharm. Sci., 72, 866 (1983).
21. M. Pikal and A. L. Lukes, J. Pharm. Sci., 65, 1269 (1976).
22. J. T. Carstensen and R. Kothari, J. Pharm. Sci., 70, 1095 (1981).
23. J. T. Carstensen, *Pharmaceutics of Solids and Solid Dosage Forms*, Wiley-Interscience, New York, 1977, pp. 6, 41.
24. J. T. Carstensen and M. Franchini, Drug Dev. Ind. Pharm., 18, 85 (1992).
25. B. Kaye, in *The Fractal Approach to Heterogeneous Chemistry* (D. Avnir, Ed.), John Wiley & Sons, Chichester, UK, 1989, p. 62.
26. J. T. Carstensen, *Pharmaceutics of Solids and Solid Dosage Forms*, John Wiley & Sons, New York, 1977, p. 56, 226, 228.
27. J. T. Carstensen and X. P. Hou, J. Pharm. Sci., 74, 466 (1985).
28. E. H. Washburn, Phys. Rev., 17, 273 (1921).
29. J. T. Carstensen, *Pharmaceutics of Solids and Solid Dosage Forms*, Wiley-Interscience, New York, 1977, pp. 11–15.
30. L. Van Campen, G. Zografi, and J. T. Carstensen, Int. J. Pharm., 5, 1 (1980).
31. L. Van Campen, G. L. Amidon, and G. Zografi, J. Pharm. Sci., 72, 1381 (1983).
32. L. Van Campen, G. L. Amidon, and G. Zografi, J. Pharm. Sci., 72, 1388 (1983).
33. L. Van Campen, G. L. Amidon, and G. Zografi, J. Pharm. Sci., 72, 1394 (1983).
34. J. T. Carstensen and F. Usui, J. Pharm. Sci., 74, 1293 (1984).

35. J. T. Carstensen, J. B. Johnson, W. Valentine, and J. Vance, J. Pharm. Sci., 53, 1050 (1964).
36. P. R. Perrier and U. W. Kesselring, J. Pharm. Sci., 72, 1072 (1983).
37. A. D. F. Toy, Inorganic phosphorous chemistry, in *Comprehensive Inorganic Chemistry* (J. C. Bailar, Jr., H. J. Emelius, R. Nyholm, and A. F. Trotman-Dickenson, Eds.), A. Wheaton & Co., Exeter, UK, 19??, pp. 389–543.
38. J. E. Bodnar, J. R. Chen, W. H. Johns, E. P. Mariani, and E. C. Shinal, J. Pharm. Sci., 72, 535 (1983).
39. J. T. Carstensen, M. Franchini, and K. Ertel, J. Pharm. Sci., 81, 303 (1992).
40. J. T. Carstensen, *Drug Stability*, Marcel Dekker, New York, 1991, p. 60.
41. L. J. Ravin, E. S. Rattie, A. Peterson, and D. E. Guttman, J. Pharm. Sci., 67, 1528 (1978).
41. Y. W. Chien, P. R. Keshary, Y. C. Huang, and P. P. Sarpotdar, Drug Dev. Ind. Pharm., 72, 968 (1983).
42. J. T. Carstensen, in *Dissolution Technology* (L. Leeson and J. T. Carstensen, Eds.), The Academy of Pharmaceutical Sciences, American Pharmaceutical Association, Washington, DC, 1974, p. 5.
43. S. Riegelman, *Dissolution Testing in Drug Development and Quality Control*, The Academy of Pharmaceutical Sciences Task Force Committee, American Pharmaceutical Association, 1979, p. 31.
44. K. Iba, E. Arakawa, T. Morris, and J. T. Carstensen, Drug Dev. Ind. Pharm., 17, 77 (1991).
45. A. Hixson and J. Crowell, Ind. Eng. Chem., 23, 923 (1931).
46. T. H. Wilson and G. Wiseman, J. Physiol., 123, 116 (1954).
47. R. K. Crane and T. H. Wilson, J. Appl. Physiol., 12, 145 (1958).
48. J. T. Dolusio, N. F. Billups, L. W. Dittert, E. J. Sugita, and J. V. Swintosky, J. Pharm. Sci., 58, 1196 (1969).
49. J. T. Carstensen, *Pharmaceutics of Solids and Solid Dosage Forms*, Wiley-Interscience, New York, 1977, pp. 99, 101.
50. J. T. Carstensen, J. L. Wright, K. W. Blessel, and J. Sheridan, J. Pharm. Sci., 67, 48 (1978).
51. J. T. Carstensen, J. L. Wright, K. W. Blessel, and J. Sheridan, J. Pharm. Sci., 67, 982 (1978).

8

Cutaneous and Transdermal Delivery: Processes and Systems of Delivery

Gordon L. Flynn
College of Pharmacy, The University of Michigan, Ann Arbor, Michigan

I. INTRODUCTION

The skin forms the body's defensive perimeter against what is in reality a hostile external environment. As such, in the normal course of living, it suffers more physical and chemical insult than any other tissue of the body. It is inadvertently scraped, abraded, scratched, bruised, cut, nicked, and burned. Insects bite it, sting it and, occasionally, furrow through it. At times it is exposed to detergents, solvents, waterborne pollutants, and myriad other chemicals and residues. Bacteria, yeasts, molds, and fungi live on its surface and within its cracks and crevices. It is brushed, smeared, dusted, sprayed, and otherwise anointed with toiletries, cosmetics, and drugs. Any of these exposures can rile the skin or provoke allergy. If there is only minor damage associated with such insults, the skin repairs itself in short order, without a trace left of the injury. If the insult is severe, its reconstruction takes far longer and may occur with scarring. Such repair is essential, for humans cannot survive for long with an extensively damaged skin. In its intact state the skin is a formidable barrier, impenetrable to otherwise life-threatening microorganisms and resistant to chemicals and tissue-harmful ultraviolet rays. The skin also keeps us from losing life's essential chemicals and fluids to the external environment. To perform these necessary functions, the skin has to be tough and, at the same time, flexible, for it is stretched and flexed continually as we move around within it. In its healthy state, it is thus a remarkable fabric, strong and far more complex than any man-made material [1].

A myriad of medicated products are applied to the skin or readily accessible mucous membranes that in some way either augment or restore a fundamental function of the skin, or pharmacologically modulate an action in the underlying tissues. Such products are referred to as *topicals* or *dermatologicals*. A *topical delivery* system is one that is applied directly to any external body surface by inunction (spreading the formula with the fingers and rubbing it in), by spraying or dusting it on, or by instilling it (as with a dropper). Thus, the term topical is frequently used in contexts for which the application is to the surface of the eye (cornea and

FRESENIUS EXHIBIT 1012
Page 187 of 293

conjunctival membranes), the external ear, the nasal mucosa or the lining of the mouth (buccal mucosa), or even the rectum, the vagina, or the urethral lining. The term, *dermatological*, on the other hand, is limited to products that are applied to the skin or the scalp and an *external use only* label is used to denote such restricted use. The distinction between general topical use and external use is not trivial. Mucous membranes offer topically applied drugs ready access to the systemic circulation, whereas normal skin is relatively impenetrable. Therefore, many drugs and chemicals can be applied to the external skin surface safely that are unsafe to place in contact with mucosal barriers. The external use only label identifies these important differences and alerts a patient to restrict the use of a product to the true skin. This chapter deals mainly with dermatologicals, but the general concepts and rationale are applicable to other modes of topical therapy as well.

The distinctions pharmacists have to make concerning topical dosage forms and their suitability for use obviously go far deeper than merely appreciating the significance of external use only labels. Pharmacological, toxicological, and risk–benefit valuations have to be made for every drug product, but the industry together with the U.S. Food and Drug Administration (FDA) has these in hand long before the pharmacist ever sees a product. Today's dosage forms, the delivery systems for the drugs, are also industry creations subject to FDA sanction. The dispensing pharmacist's lack of input into drug development does not abrogate his or her responsibility for assuring that dispensed products conform to high standards, however. Consequently, the consummate professional takes every opportunity to evaluate, by literature or by observation, how the performances of various dosage forms measure up to absolute standards and stack up against one another. A pharmacist should thus be seeking answers to the following questions: Is the drug bioavailable as administered? Are the drug and the dosage form stable? Is the formulation free of contamination? Is it pharmaceutically elegant? The particulars relative to these attributes depend on the type of dosage form and the route of administration. Regardless, a *no* to any of the first three questions is reason to remove a product from distribution. Elegance may be sacrificed some for function, but only in degree. One of the goals here is to explicate the performance attributes of dermatological dosages forms to use as a basis for product evaluation.

II. THE STRUCTURE AND FUNCTION OF SKIN

To answer questions about the therapeutic and cosmetic uses of the myriad dermatological concoctions that are available, a pharmacist must be knowledgeable about the anatomical structure and physiological functions of the skin and the chemical compositions and physicochemical properties of its constituent phases. Furthermore, some understanding of how the latter attributes are affected by disease and damage is a must, as is knowledge of how the skin's physiology and function vary with age, environmental conditions, and other factors. Rational approaches to topical therapy rest on such insight.

A. Skin Functions

General functions of the skin are outlined in Table 1. These functions include containment of tissues and organs, multifaceted types of protection, environmental sensing, and body heat regulation. Some skin functions are inextricably entwined. For instance, containment and the barrier functions are to some extent inseparable. Active sweating is accompanied by increased peripheral blood flow which, in turn, is tied in with vascular nourishment of the cells of the skin as needed to promote their proliferation, differentiation, and specialization.

FRESENIUS EXHIBIT 1012
Page 188 of 293

Table 1 Functions of the Skin

Function
Containment of body fluids and tissues
Protection from harmful external stimuli (barrier functions)
Microbial barrier
Chemical barrier
Radiation barrier
Thermal barrier
Electrical barrier
Reception of external stimuli
Tactile (pressure)
Pain
Thermal
Regulation of body temperature
Synthesis and metabolism
Disposal of biochemical wastes (through secretions)
Intraspecies identification and or attraction (apocrine secretions)
Blood pressure regulation

Source: Refs. 2 and 3.

Let us consider how the skin is structured so that we might better understand how it performs some of its vital functions, beginning with the cross section of the skin sketched in Fig. 1. This illustration shows the readily distinguishable layers of the skin are, from the outside of the skin inward: (a) the ≈ 10-μm–thin, devitalized outer epidermis, called the stratum corneum; (b) the ≈ 100-μm–thin living epidermis; and (c) the ≈1000-μm–thin (1-mm-thin) dermis, all these stated thicknesses being representative only, for the actual thicknesses of these strata vary severalfold from place to place over the body. Dispersed throughout the skin, in various numbers and size, depending on body site, one finds several skin glands and appendages, namely (a) hair follicles and their associated sebaceous glands (pilosebaceous glands); (b) eccrine sweat glands; (c) apocrine sweat glands; and (d) nails (finger and toe). Each has unique population densities and distributions at disparate body locations. There are characteristic differences in appearances of the structures from place to place on the body as well.

A highly complex network of arteries, arterioles, and capillaries penetrates the dermis from below and extends up to the surface of, but not actually into, the epidermis. A matching venous system siphons the blood and returns it to the central circulation. The flow of blood through this vasculature is integrated with the production and movement of lymph through the dermal lymphatics. The dermis is laced with tactile, thermal, and pain sensors.

B. Stratum Corneum

The outermost layer of the skin appearing in the exploded epidermal sketch of Fig. 1 represents the stratum corneum or *horny layer* of the skin; the main element of the skin's permeation barrier, it is a multicellular, essentially metabolically inactive tissue comprised of acutely flattened, stacked, hexagonal cell building blocks formed from once-living cells. These cellular building blocks are layered 15–25 cells deep over most of the body [2]. Sometimes the cells appear stacked one on top of the next in neat columns, but on most occasions they are irregularly arranged. The stratum corneum exhibits regional differences in thickness, being as thick as several hundred micrometers on the friction surfaces of the body such as the palms and soles. However, over most of the body the thickness is approximately 10 μm, less than a fifth the thickness of an ordinary piece of paper [2,4]. It is a dense tissue, about 1.4 g/cm^3 in the

FRESENIUS EXHIBIT 1012
Page 189 of 293

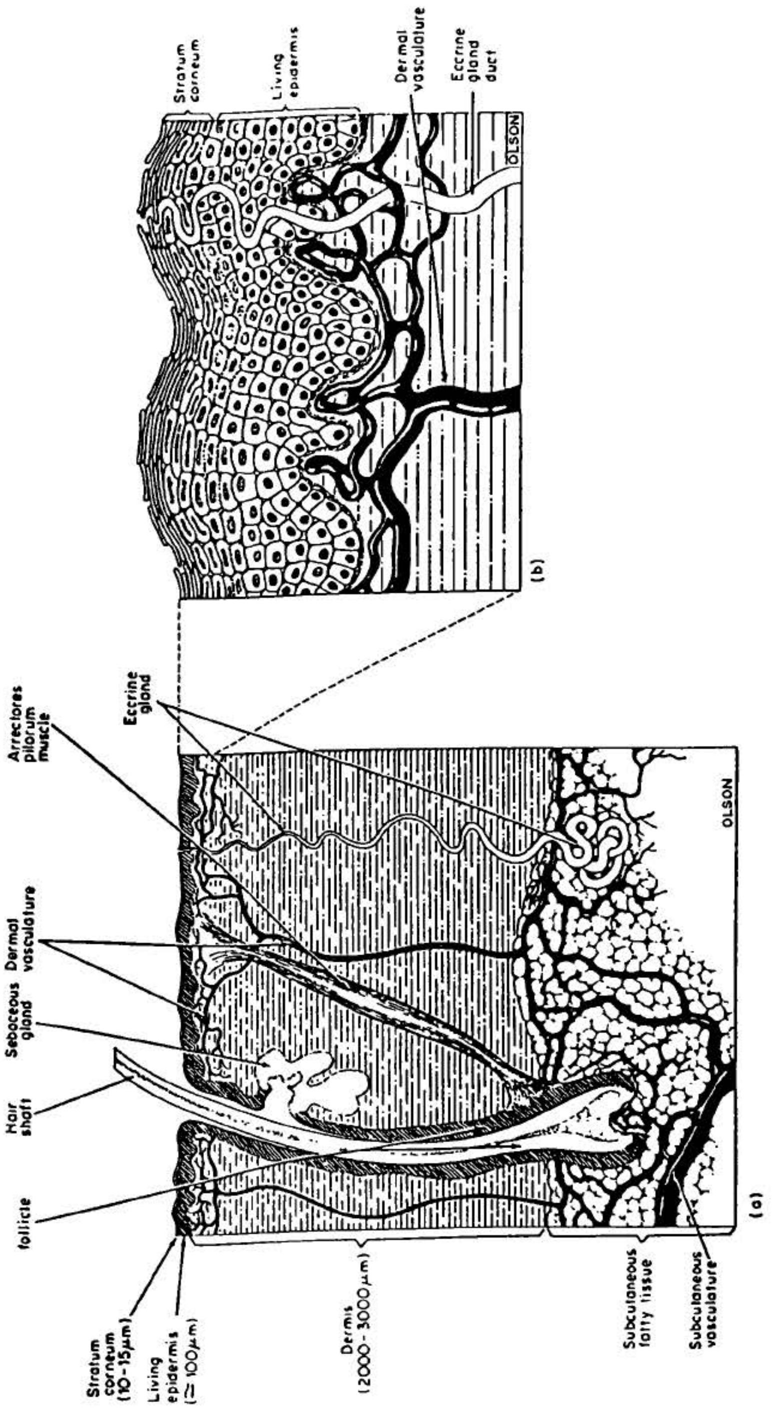


Fig. 1 Sketch of the skin.

FRESENIUS EXHIBIT 1012
Page 190 of 293

dry state. Consequently, the stratum corneum is also occasionally referred to as the stratum compactum.

The stratum corneum is under continuous formation. Cells that are worn off the surface are replaced from beneath, one for one, with complete turnover of the layer occurring every 2 weeks in normal individuals [5]. In humans the cells that give rise to the stratum corneum originate exclusively in the basal (basement) layer of the epidermis. Thus, this layer is often referred to as the germinative (proliferative) layer. Mitosis begins an extraordinary process in which daughter cells are pushed outward, first to form the so-called spinous or prickle cell layer, and then serially, the granular, lucid, and horny layers. During their transit through the epidermal mass, the cells flatten acutely and internally synthesize the protein and lipid components that eventually characterize the fully differentiated horny layer. Individual strands of structural protein are formed, even while the cells reside in the basal layer. As the cells progress upward through the epidermis to take positions in the stratum corneum, their protein content expands, so much so that massed proteins of several kinds are distinguishable by the time the cells reach the granular layer. A basic protein here, which stains deeply to give the granular layer its characteristic histological appearance, is filagrin. Filagrin (*fil*ament *ag*gregating pro*tein*), which is released in one of the culminating events of formation of the stratum corneum, induces individual helical strands of prekeratin protein to twist together into multistranded fibers that themselves have helical geometry. These fibers, in turn, are spontaneously bundled and concentrated so that the intracellular space of the fully differentiated horny cell is literally packed full with this semicrystalline α-keratin and its amorphous keratin counterpart (β-keratin). Indeed, nothing else is seen there in the electron microscope. The intracellular space, therefore, is dense and, to a great degree, chemically impenetrable.

Lipid is also synthesized during a keratinocyte's epidermal transit and is collected in small vesicles visible in the granular layer. These were designated membrane-coating granules by microscopists long before the their content and function were known. As the granular cells transform, these vesicles migrate to the upper cell membrane, at which point their contents are passed exocytotically into the intercellular space. This lipid functionally becomes a mortar that seals the horny structure, making the stratum corneum, pound for pound, an incredibly efficient moisture barrier. Virtually all the lipid of the stratum corneum is in this interstitial space, much of it being there in liquid crystalline, bilayer arrays [6]. An exoskeleton (infrastructure) consisting of residual cell membranes, bound together by desmosomes and tonofibrils, separates these keratin and lipid domains. The lipid content of the horny layer represents about 20% of the stratum corneum's dry weight, whereas the endoskeleton amounts to roughly 5%, as indicated in Table 2 [2,5,7].

In its normal state at ordinary relative humidities, the stratum corneum takes up moisture to the extent of 15–20% of its dry weight [2]. However, on some areas of the body, this water content can increase to several multiples of the dry weight when the skin is waterlogged. The stratum corneum also becomes highly hydrated when natural evaporation of water from the skin's surface, so-called insensible perspiration, is blocked by an occlusive dressing. Regardless of how its water content is increased, this tissue becomes more pliable and molecules diffuse through it with greater facility. It is likely that some substances become more soluble in it as well. Conversely, as the stratum corneum dries out, it becomes brittle. Ultradry, inelastic horny tissue tends to split and fissure when stretched, giving rise to conditions we know as chapped lips, windburn, and dishpan hands.

The stratum corneum is thus a dense, polyphasic, epidermal sheathing made from dehydrated and internally filamented former cells held together by desmosomes, tonofibrils (intercellular anchors), and interstitial lipid. It has been estimated that the stratum corneum contains ten times the fibrous material of the living epidermis in roughly one-tenth the space [8]. At its

FRESENIUS EXHIBIT 1012
Page 191 of 293

Table 2 Composition of the Stratum Corneum

Tissue component	Gross composition	Percentage of dry weight
Cell membrane	Lipid, protein	≈ 5
Intercellular space	Mostly lipid, some protein and polysaccharide	≈ 20
Intracellular space	Fibrous protein (≈ 65–70%), non-fibrous (soluble) protein (≈ 5–10%)	≈ 75
Overall protein	Water soluble (10%), keratin (≈ 65%), cell wall (≈ 5%)	70–80
Overall lipid		10–20
All other		Up to 10
Water (normal hydration)		15–20
Water (fully hydrated)		Upwards of 300

Source: Data from Refs. 2 and 4

undersurface the stratum corneum is in contact with the living epidermal mass, and at its other surface with the environment. The underlying cells contain water with the high thermodynamic activity of the physiological milieu. As a rule, air at the surface of the skin is dry and of low water activity. Consequently, water diffuses outwardly (*down* this activity gradient), eventually escaping into the environment, a process known as *insensible perspiration*. Over normal skin about 5 ml of water is lost this way per square meter of body surface per hour (or 0.5 mg cm^{-2} hr^{-1}) [9]. Since an adult has upward of 2 m^2 of body surface area, whereas an infant has roughly 0.25 m^2 and a 2-year-old 0.75 m^2 or so, an adult loses about 250 ml (8 oz) of water per day by this mechanism. A small child loses less than half this amount. Such water loss can increase alarmingly and to as much or more than 100 ml hr^{-1} m^{-2} over skin ravaged by disease or damage. High evaporation rates such as this, over expansive areas, carry away sufficient body heat to lower the core temperature to dangerous levels.

C. Viable Epidermis

The animate cells of the epidermis have a readily definable upper interface with the lifeless stratum corneum and an even more clearly demarcated, deep interface with the dermis (see Fig. 1). In drug delivery considerations, the whole of this tissue is considered as a singular diffusional field, although, when viewed under microscope, it is clearly multilayered. The identifiable strata are, from bottom to top, the basal layer [stratum (s.) germinativum], a monolayer of cubical or columnar cells, otherwise unremarkable in appearance; the spinous or prickle layer (s. spinosum), where cells exhibit sharp surface protuberances; the granular layer (s. granulosum), which takes up stain to yield a mottled appearance; and, in some histological displays, the lucid layer (s. lucidum), which appears translucent. These layers reflect a progressive differentiation in the cells, which eventuates in their death and placement within the horny structure. When physicochemically considered, the viable epidermis is nothing more than a wedge of tightly massed cells, literally small compartments of cytoplasm encapsulated within delicate cell membranes, held together by tonofibrils. Here, as elsewhere in the body, water has an activity equivalent to that of a highly dilute NaCl solution (0.9% NaCl). Relative to diffusion, the density and consistency of this cell composite are only a little greater than found with water.

The viable epidermis makes a flat interface with the stratum corneum. Its interface with the dermis is papillose (mounded). Myriad tiny bulges of the epidermis fit with exacting reciprocity

FRESENIUS EXHIBIT 1012
Page 192 of 293

over dermal depressions and ridges. It is these ridges that give the friction surfaces of the body their distinctive patterns (e.g., fingerprints). And, since hair follicles and eccrine glands have epidermal origins, epidermal cells actually extend through the dermis and into subcutaneous tissue by way of these tiny glands (see Fig. 1). This has a generally unappreciated significance in terms of the self-repair capabilities of the skin. As long as an injury does not extend to the base of the glands, islands of vital cells capable of regenerating a scar-free skin surface remain available for repair. Discounting these deep rootages, the epidermis is on the order of 100 μm thick [10].

The principal cells of the epidermis are its keratinocytes. One also finds Langerhans cells—cells of white blood cell progeny—at regular intervals within the bulk of the epidermal mass, and melanocytes strategically placed in the basal layer just above the epidermal–dermal junction (Table 3). Langerhans cells function as antigen-presenting cells in the skin's immunological responses. Under the influence of melanocyte-stimulating hormone (MHS), melanocytes synthesize the pigment that gives the races of humans their unique skin colorations. Melanocytes are also set into action by ultraviolet radiation, leading to suntanning. Other cells, occasionally seen in skin sections, are migrant macrophages and lymphocytes. These are particularly numerous when the skin is traumatized.

D. Dermis

The dermis, as depicted in Fig. 1, is a nondescript region lying between the epidermis and the subcutaneous, fatty region. In reality it is a complex structure, consisting mainly of a meshwork of structural fibers, collagen, reticulum, and elastin, filled with a mucopolysaccharidic gel, called the ground substance [2]. Approximate proportions of these phases are indicated in Table 4. The dermis ranges from 1 mm (1000 μm) to 5 mm in thickness [11]. The upper one-fifth or so of the tissue, the papillary layer by name, is finely structured and is the support for the delicate capillary plexus that nurtures the epidermis. The papillary dermis eventually merges with the far coarser fibrous matrix of the reticular dermis. This deeper layer is the main structural element of the dermis and, for that matter, of the skin. Of considerable importance, the microcirculation that subserves the skin is entirely located in the dermis. The dermis is also penetrated by a network of sensory nerves (pressure, temperature, and pain) and a rich lymphatic network. Numerous fibroblasts, cells that synthesize the structural fibers, are found here [2], and one also finds mast cells scattered about (see Table 3). The latter are thought to play

Table 3 Cells of the Skin

Cell type	Principal function
Cells of the epidermis	
Keratinocytes	Form keratinized structures
Langerhans cells	Antigen presentation
Melanocytes	Pigment synthesis
Macrophages, lymphocytes	Migrant cells, immune responses
Cells of the dermis	
Fibroblasts	Fiber synthesis
Mast cells	Make ground substance, histamine
Blood cells	
Endothelial cells	Form the blood vessels
Nerve cells and endings	Sensors

Source: Data from Refs. 2 and 11.

Table 4 Composition of the Dermis

Component	Approximate % composition
Collagen	75.0
Elastin	4.0
Reticulin	0.4
Ground substance	20.0

Source: Data from Refs. 2 and 11.

a role in synthesizing ground substance and are known to be a source of the histamine that is released when the skin is immunologically provoked.

E. Skin's Circulatory System

Arteries entering the skin arise from more substantial vessels located in the subcutaneous connective tissue. These offshoots form a plexus just beneath the dermis [11]. Branches from this subcutaneous network supply blood directly to the hair follicles, the glandular appendages, and the subcutaneous fat. Branches to the upper skin from this deep plexus divide again within the lower dermis, forming a deep subpapillary network. Arterioles reaching the upper dermis out of this more distal plexus are on the order of 50 μm in diameter and exhibit arteriovenous anastomoses, shuntlike connections that link the arterioles directly to corresponding venules. The dermal arterioles then branch to form a shallower, subpapillary plexus that supplies twigs to the dermis and fine-branched capillary loops to the papillae at the dermal–epidermal interface. The epidermis itself is avascular.

The veins of skin are organized along the same lines as the arteries in that there are both subpapillary and subdermal plexuses [11]. The main arteriole communication to these is the capillary bed. Copious blood is passed through capillaries when the core body is either feverish or overheated, far more than needed to sustain the life-force of the epidermis, and this rich perfusion lends a red coloration to skin. When there is the opposite physiological need, the capillary bed is short-circuited, and blood is passed directly into the venous drainage by way of the arteriovenous anastomoses. Fair skin blanches when this occurs. These mechanisms act, in part, to regulate body temperature and blood pressure.

The vascular surface available for exchange of substances between the blood and local tissue has been estimated to be of the same magnitude as that of the skin (i.e., 1–2 cm^2/cm^2 of skin). At room temperature, about 0.05 ml of blood flows into the skin per minute per gram of the tissue; the supply increases considerably when the skin is warmer [4,12]. Ordinarily, sufficient blood reaches to within 150 μm or so of the skin's surface to efficiently draw chemicals into the body that have percutaneously gained access to this depth [7]. Interestingly, this local circulation is turned down by vasoconstrictors (e.g., glucocorticoids) and up by vasodilators (e.g., nicotine), respectively. Such responses are so reliable that vasoconstriction (blanching) has become an FDA-sanctioned index of corticosteroid penetration of the skin [13,14]. The relation between blood flow and local clearance of percutaneously absorbed drugs, including aspects of vasoconstriction and vasodilation, is not well known.

Lymphatics of the skin extend up and into the papillary layers of the dermis. A dense, flat meshwork of lymphatic capillaries is found here [11]. These lymphatic vessels pass to a deeper network at the lower boundary of the dermis. Serum, macrophages, and lymphocytes easily pass through the interfaces of the skin's lymphatic and vascular networks.

FRESENIUS EXHIBIT 1012
Page 194 of 293

F. Skin Appendages

Hair follicles and their associated sebaceous glands (pilosebaceous glands), eccrine glands, apocrine glands, and nail plates are referred to as the skin's appendages. Hair follicles are found within the skin everywhere except the soles, the palms, the red portion of the lips and the external genitalia. They are formed from epidermal cells in fetal life. From place to place, the follicles and the hair they produce differ markedly in prominence. Delicate primary hair is found on the fetus; secondary hair or *down* covers the adult forehead; terminal hair ordinarily blankets the scalp, and it covers the pubic region and underarms [2]. A hair (hair shaft) emerges from each *follicle*, as shown in Fig. 1. The follicle itself lies within the skin and consists of concentric layers of cellular and noncellular components positioned in the skin at a slight angle. Each follicle is anchored to the surrounding connective tissue by an individual strand of smooth muscle, the arrector pilorum, contraction of which causes the hair to stand upright, raising goose pimples on human skin. In animals like cats, hair stands on end as part of the flight-or-flight response.

The hair shaft is formed continuously by cell division, differentiation, and compaction within the bulb (base) of each active hair follicle, a process that is completed deep in the follicle. Hair, like stratum corneum, is thus a compact of fused, keratinized cells. Collectively, hair follicles occupy about 1/1000 of the skin's surface [10,15], a factor that sets a limit on the role they can play as a route of penetration. Each hair follicle possesses one or more flasklike sebaceous glands (see Fig. 1). These have ducts that vent into the open space surrounding the hair shaft just below the skin's surface. Just as with keratinocytes, the cells of sebaceous glands, *sebocytes*, are programmed to differentiate and die. Before they die and disintegrate, they pack themselves full of lipid-containing vesicles. The residue left behind at their death is mixed with other follicular debris deep within the follicular orifice to form the actual substance, sebum, expressed onto the skin. Sebum is then forced upward around the hair shaft and onto the skin's surface through outlets having diameters ranging from 200 to 2000 μm (2 mm), depending on body location [11]. Glands with the largest openings are found on the forehead, face, nose, and upper back. These contain only a tiny hair if they contain one at all.

Eccrine or salty sweat glands are found over the entire body, except the genitalia. Of fetal epidermal origin, they consist of tubes extending from the skin surface to the footings of the dermis. Here the tube coils into a ball roughly 100 μm in diameter (see Fig. 1) [11]. By anatomical count, there are between 150 and 600 glands per cubic centimeter of body surface, depending on the site [16]. They are particularly concentrated in the palms and soles, attaining a densities in these locations well in excess of 400 glands per centimeter. However, since many of the glands remain dormant, estimates of their numbers are appreciably lower if based on actual sweating. Each gland has a microscopic orifice within the surface of the skin of about 20 μm diameter from which its secretions are spilled. In total, these glandular openings represent approximately 1/10,000 of the skin's surface [10]. Eccrine sweat is a dilute (hypotonic), slightly acidic (pH ≈ 5.0 owing to traces of lactic acid) aqueous solution of salt. Its secretion is stimulated when the body becomes overheated through warm temperatures or exercise. Evaporation of the water of the sweat cools the body's surface and, thereby, the body. Since the gland is innervated by the autonomic nervous system, eccrine sweating is also stimulated emotionally (the clammy handshake).

Apocrine glands have highly regionalize locations and are found only in the axillae (armpits), in anogenital regions, and around the nipples. Along with other secondary sexual characteristics, the glands develop at puberty. We know that they are innervated emotionally and through concupiscence. In the mature female, they exhibit cyclic activities in harmony with the menstrual cycle. Similar to eccrine glands, they are coiled tubular structures, but they are

FRESENIUS EXHIBIT 1012
Page 195 of 293

roughly ten times larger. Therefore, they extend entirely through the dermis and well into the subcutaneous layer [2,11]. Each gland is paired up with a neighboring hair follicle, and its secretion is vented into the sebaceous duct of the follicle beneath the surface of the skin. Because the secretion of the apocrine gland is combined with sebum before reaching the skin's surface, its chemical makeup is a enigma. What is not a mystery is that bacterial decomposition of the secretion is responsible for human body odor.

III. SKIN FUNCTIONS

The skin's main physiological roles are outlined in Table 1. Of these, the chemical barrier function is central to the use of topical drugs because deposition of a topical drug into the deeper, living strata of skin is a prerequisite for achieving its pharmacological effect. Degeneration in some of the functions can be pathognomonic of disease. Even when specific functions do not relate materially to the skin's state of health, they are tied in with cosmetic practices and, thereby, are of interest to pharmacists.

A. Containment

The containment function relates specifically to the ability of the skin to confine underlying tissues and restrain their movements. The skin draws the strength it needs to perform this mechanical role from its tough, fibrous dermis [2]. Ordinarily, the skin is taut, even when under resting tension, yet it stretches easily and elastically when the body is in motion, quickly returning to normal contours when the stretching ceases. This extensibility of the skin is attributable to an alignment of collagen fibers, under tension and in the direction of a load, that are otherwise nonaligned in the ground tension state. Elastin fibers attached to individual collagen strands relax and, in doing so, restore the irregular order of the restive state. As one ages, the resilience of these dermal fibers decreases and the tensile strength of the tissue increases. Eventually, the skin becomes stretched beyond its ability to elastically restore its initial condition and it folds over itself or *wrinkles*. Lost elasticity is advanced through extended exposure to ultraviolet radiation (sunlight); thus wrinkling is often pronounced on dedicated sunbathers.

The behavior of the epidermis when distended is also of importance. Obviously, this layer should not be torn or broken when placed under mechanical stress, for an intact epidermis is the body's first line of defense against infection. It is the stratum corneum's role to fend against tearing [2]. Pound for pound, this tissue is actually stronger than the dermal fabric and, as a rule, it is sufficiently elastic to adjust to stretching. Its pliability, however, is conditional, and it fissures and cracks if stretched when excessively dry. Arid atmospheres alone can produce this condition (windburn). Detergents and solvents, which extract essential, water-sequestering lipids from the stratum corneum, and diseases, such as psoriasis, that are associated with a malformed horny structure, render the stratum corneum brittle and prone to fissuring.

Although much is still to be learned about the factors that contribute to the pliability of the stratum corneum, it is generally accepted that its elasticity is dependent on a proper balance of lipids, hygroscopic, water-soluble substances, and water, all in conjunction with its keratin proteins. Water is its principal plasticizer, or softening agent, and it takes roughly 15% moisture to maintain adequate pliability. The capacity of the stratum corneum to bind and hold onto water is greatly reduced by extracting it with lipid solvents such as ether and chloroform. Moreover, there is a further significant decrease in the water-binding capacity of callus, a thickened stratum corneum found on the palms and soles, when it is extracted with water after having first been treated with a lipid solvent. The latter observation seems to tell us that amino

acids, hydroxy acids, urea, and inorganic ions, cosmetically referred to at the skin's *natural moistening factor*, and the stratum corneum's lipids both assist the stratum corneum in retaining moisture necessary to plasticize its mosaic, filamented matrix. In effect the water makes the tissue less crystalline through its interposition between polymer strands.

B. Microbial Barrier

Normal stratum corneum, taken in its entirety, is a dense, molecular continuum penetrable only by molecular diffusion. It is virtually an absolute barrier to microbes, preventing them from reaching the viable tissues and an environment suitable for their growth. The outermost stratum corneum is continuously being shed in the form of microscopic scales (natural desquamation) and, to a limited depth, is laced with tiny crevices. Many microorganisms—pathogens and harmless forms alike—are found in these rifts. Surgeons know well that superficial washing is insufficient to remove these surface microbes; therefore, the surgical scrub is an energetic and intense cleansing with a disinfectant soap. The microorganisms residing on and in the skin can and do initiate infections if seeded into living tissues as a result of abrasive or disease-induced stratum corneum damage. Consequently, antiseptics and antibiotics are widely used to chemically sanitize wounds.

Beyond physical barrier protection, several natural processes lead to skin surface conditions unfavorable to microbial growth. Both sebaceous and eccrine secretions are acidic, lowering the surface pH of the skin below that welcomed by most pathogens. This *acid mantle* (pH $\approx$ 5) [16] is moderately bacteriostatic. Sebum also contains a number of short-chain fungistatic and bacteriostatic fatty acids, including propanoic, butanoic, hexanoic, and hepatanoic acids [17]. That the skin's surface is dry also offers a level of protection. It comes as no surprise that fungal infections and other skin infections are more prevalent in the skin's folds during warm weather, as intensified sweating leaves the skin continually moist in these regions.

Glandular orifices provide possible entry points for microbes. The duct of the eccrine sweat gland is tiny and generally evacuated. Experience tells us that this is not an easy portal of entry, although localized infection is seen occasionally in infants suffering prickly heat. Pilosebaceous glands seem more susceptible to infection, particularly those on the forehead, face, and upper back, referred to as sebaceous follicles. Glands at these specific locations have an almost imperceptible hair surrounded by a massive sebaceous apparatus and are especially prone to occlusion and subsequent infection (acneform pimples and blackheads). Such sebaceous gland infections are usually localized. However, if the infected gland ruptures and spews it contents internally, deep infection is possible. The body defends against this by walling off the lesion (forming a sac or *cyst*) and then destroying the eliminating the infected tissue. The destruction caused by cystic acne is deep, so much so that facial scarring is associated with it. In hair follicles containing prominent hairs, the growing hair shaft acts as a sebum conveyer that unblocks the orifice. It may be strictly coincidental, but such follicles seem less prone to clogging and infection.

C. Chemical Barrier

The intact stratum corneum also acts as a barrier to chemicals brought into contact with it. Its diffusional resistance is orders of magnitude greater than found in other barrier membranes of the body. Externally contacted chemicals can, in principle, bypass the stratum corneum by diffusing through the ducts of the appendages. The ability of each chemical to breach the skin and the diffusional route or routes it takes are dependent on its own physicochemical properties and the interactions it has within the skin's various conduit regimes. Being central to the

effectiveness of dermatological products, exposition of the skin's barrier properties is made in a following section.

D. Radiation Barrier

Ultraviolet wavelengths of 290–310 nm from the UV-B band of radiation constitute the principal tissue-damaging rays of the sun that are not fully atmospherically filtered. One hour of exposure to the summer sun and its damaging rays can produce a painful burn, with a characteristic erythema. The skin has natural mechanisms to prevent or minimize such sun-induced trauma, but it takes time to set these into place. On stimulation by ultraviolet rays, particularly longer, lower-energy rays above 320 nm, melanocytes at the epidermal–dermal junction produce the pigment, melanin. Melanin's synthesis begins in the corpus of the melanocyte, with pigment-forming granules migrating outward to the tips of the long protrusions of these starlike cells. Adjacent epidermal cells endocytotically engulf these projections. Through this cellular cooperation, melanin, which absorbs and diffracts harmful UV rays, becomes dispersed throughout the epidermis, and a person tans, with his or her capacity to sunburn declining accordingly. It should be realized that tanning takes time, several days in fact, and is incapable of protecting a person on first exposure. Damaging ultraviolet exposure also stimulates epidermal cell division and thickening of the epidermis (acanthosis). Such thickening, too, takes several days. When effected, it also lends protection to the underlying tissues.

Pharmacists should tell their sun-deprived, fair-skinned patrons not to spend more than 15–20 min in the midday sun (10:00 a.m. to 3:00 p.m.) on first exposure when traveling to vacation spots such as Florida [18]. This is ample, safe exposure to initiate the tanning response in those who are able to tan. Exposures can be increased incrementally 15 min/day until a 45-min tolerance is developed, which is generally an adequate level of sun protection in conjunction with the use of sunscreens. It should be obvious that dark-skinned people are already heavily pigmented and, thus, far less susceptible to burning. Other individuals do not tan at all and must apply sunscreens with high protection factors before sunbathing.

E. Electrical Barrier

Dry skin offers a high impedance to the flow of an electrical current [3]. Stripping the skin, by successively removing layers of the stratum corneum with an adhesive tape, reduces the electrical resistance about sixfold, which tells us that the horny layer is the skin's prime electrical insulator. Its high impedance complicates the measurement of body potentials, as is done in electroencephalograms and electrocardiograms. Consequently, electrodes having large contact areas are used to monitor the brain's and the heart's electrical rhythms. Granular salt suspensions or creams and pastes containing high percentages of electrolytes are placed between the electrode surface and the skin to assure the electrical conductance is adequate to make the measurements.

F. Thermal Barrier and Body Temperature Regulation

The body is basically an isothermal system fine-tuned to 37°C (98.6°F). The skin has a major responsibility in temperature maintenance. When the body is exposed to chilling temperatures that remove heat faster than the body's metabolic output can replace it, changes take place in the skin to conserve heat. Conversely, when the body becomes overheated, physiological processes come into play that lead to cooling.

The skin's mechanism of heat conservation involves its very complex circulatory system [2,3]. To conserve heat, blood is diverted away from the skin's periphery by way of the

arteriovenous anastomoses. The blood's external-most circulation is effectively shut down, leading to a characteristic blanching of the skin in fair-skinned individuals. Less heat is irradiated and convectively passed into the atmosphere. Furry mammals have yet another mechanism to conserve body heat. Each tiny arrector pilorum stands its hair up straight, adding appreciable thickness to the insulating air layer entrapped in the fur, reducing heat loss.

When the body is faced with the need to cast out thermal energy, the circulatory processes are reversed, and blood is sent coursing through the skin's periphery, maximizing radiative and convective heat losses. This process produces a reddening in light skin, a phenomenon that is particularly noticeable following strenuous exercise. Exercise also leads to profuse eccrine sweating, a process that is even more efficient in heat removal. Watery sweat evaporates, with the heat attending this process (heat of vaporization) cooling the skin's surface. Factors that accelerate the evaporative process, such as the gentle flow of air produced by a fan, accelerate cooling. Low humidity favors evaporation, and one is more comfortable and sweating is less noticeable when the air's moisture content is low. Pharmacist's should be aware that eccrine sweating is a vital process not to be tampered with. Coverage of the body with a water-impermeable wrapping, as has occasionally been done in faddish weight control programs, may result in hyperthermia, particularly if there is concurrent exercise. In its extreme, hyperthermia can be fatal.

IV. RATIONALE FOR TOPICALS

One's grasp of topical dosage forms and their functioning can be nicely organized into several broad usage categories. For instance, many products exist to augment the skin barrier (Table 5). Sunscreens and anti-infectives obviously do this. The barrier is made pliable and restored in function by emollients. Pastes are sometimes used to directly block out sunlight and, at other times, to sequester irritating chemicals that would otherwise penetrate into the skin. Even insect repellants add function to the barrier.

A second general purpose of topical application involves the selective access drugs have to epidermal and dermal tissues when administered this way. Penetration of the skin can drench the local tissues with the drug before its systemic dissemination and dilution. As a result, the drug's systemic levels are kept low and pharmacologically inconsequential. In contrast, systemic treatment of local conditions bathes highly blood-perfused tissues with the drug first, with the drug's systemic effects or its side effects sometimes overpowering the actions sought for it in the skin.

In a few instances, drugs are applied to the skin to actually elicit their systemic effects. This is called transdermal therapy. Transdermal therapy is set apart from local treatment on several counts. It is only possible with potent drugs that are also highly skin permeable. To be used transdermally, compounds must be free of untoward cutaneous actions as well. When these

Table 5 Barrier Augmentation by Topical Products

Product type	Barrier effect
Sunscreens	Enhance radiation barrier
Topical anti-infectives	Augment microbial barrier
Emollients	Moisturize stratum corneum, restore barrier
Insect repellents	Add a chemical barrier to insects
Poison ivy products	Negate antigens, augment chemical barrier
Diaper rash products	Build up moisture and chemical barrier

FRESENIUS EXHIBIT 1012
Page 199 of 293