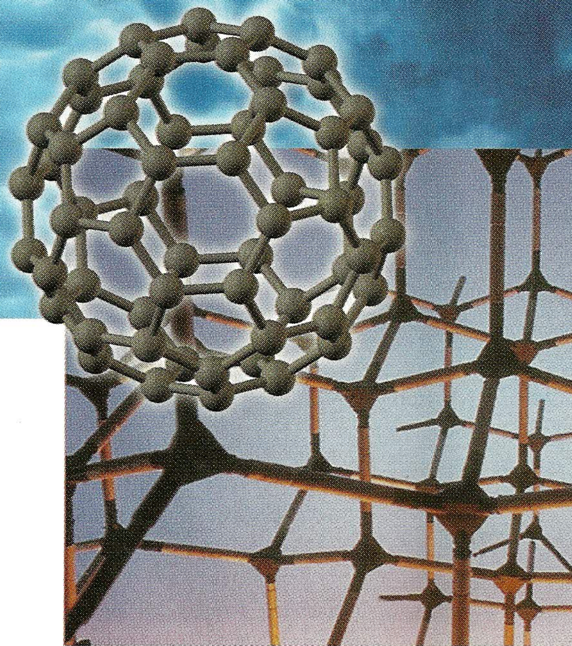


Organic Chemistry

Fourth Edition

William H. Brown

Beloit College

Christopher S. Foote

University of California, Los Angeles

Brent L. Iverson

University of Texas, Austin

THOMSON

BROOKS/COLE

Australia • Canada • Mexico • Singapore • Spain
United Kingdom • United States

IPR2021-00880, -00881
Ex. 3002

THOMSON
BROOKS/COLE™

Publisher, Physical Sciences: David Harris
Development Editor: Sandra Kiselica
Assistant Editors: Alyssa White and Annie Mac
Editorial Assistant: Candace Lum
Technology Project Manager: Donna Kelley
Marketing Manager: Amee Mosley
Marketing Assistant: Michele Colella
Advertising Project Manager: Nathaniel Bergeson-Michelson
Project Manager, Editorial Production: Lisa Weber
Creative Director: Rob Hugel
Print Buyer: Barbara Britton
Permissions Editor: Kiely Sexton

COPYRIGHT © 2005 Brooks/Cole, a division of Thomson Learning, Inc. Thomson Learning™ is a trademark used herein under license.

ALL RIGHTS RESERVED. No part of this work covered by the copyright hereon may be reproduced or used in any form or by any means—graphic, electronic, or mechanical, including but not limited to photocopying, recording, taping, Web distribution, information networks, or information storage and retrieval systems—without the written permission of the publisher.

Printed in the United States of America
3 4 5 6 7 08 07 06

For more information about our products, contact us at:
Thomson Learning Academic Resource Center
1-800-423-0563
For permission to use material from this text or product, submit a request online at **http://www.thomsonrights.com**
Any additional questions about permissions can be submitted by email to **thomsonrights@thomson.com**

COPYRIGHT © 2005 Thomson Learning, Inc. All Rights Reserved. Thomson Learning WebTutor™ is a trademark of Thomson Learning, Inc.

Library of Congress Control Number: 2004110543

Student Edition: ISBN 0-534-46773-3

Instructor's Edition: ISBN 0-534-46774-1

International Student Edition: ISBN 0-534-39597-X (Not for sale in the United States)

Production Service: Progressive Publishing Alternatives
Text Designer: Tani Hasegawa
Photo Researcher: Dena Digilio-Betz
Copy Editor: Sara Black
Illustrators: J/B Woolsey Associates and Progressive Information Technologies
Cover Designer: Didona Design
Cover Images: Molecules: Brent Iverson; blue images: Siede Preis/Getty Images and Joe Ginsberg/Getty Images
Cover Printer: Phoenix Color Corp
Compositor: Progressive Information Technologies
Printer: Courier Corporation, Kendallville

Thomson Brooks/Cole
10 Davis Drive
Belmont, CA 94002-3098
USA

Asia
Thomson Learning
5 Shenton Way #01-01
UIC Building
Singapore 068808

Australia/New Zealand
Thomson Learning
102 Dodds Street
Southbank, Victoria 3006
Australia

Canada
Nelson
1120 Birchmount Road
Toronto, Ontario M1K 5G4
Canada

Europe/Middle East/Africa
Thomson Learning
High Holborn House
50/51 Bedford Row
London WC1R 4LR
United Kingdom

Latin America
Thomson Learning
Seneca, 53
Colonia Polanco
11560 Mexico D.F.
Mexico

Solution

The isoelectric point of histidine is 7.64. At this pH, histidine has a net charge of zero and does not move from the origin. The pI of cysteine is 5.02; at pH 7.64 (more basic than its isoelectric point), cysteine has a net negative charge and moves toward the positive electrode. The pI of lysine is 9.74; at pH 7.64 (more acidic than its isoelectric point), lysine has a net positive charge and moves toward the negative electrode.

Problem 27.3

Describe the behavior of a mixture of glutamic acid, arginine, and valine on paper electrophoresis at pH 6.0.

27.3 Polypeptides and Proteins

Peptide bond The special name given to the amide bond formed between the α-amino group of one amino acid and the α-carboxyl group of another amino acid.

Dipeptide A molecule containing two amino acid units joined by a peptide bond.

Tripeptide A molecule containing three amino acid units, each joined to the next by a peptide bond.

Polypeptide A macromolecule containing many amino acid units, each joined to the next by a peptide bond.

In 1902, Emil Fischer proposed that proteins are long chains of amino acids joined together by amide bonds between the α-carboxyl group of one amino acid and the α-amino group of another. For these amide bonds, Fischer proposed the special name **peptide bond.** Figure 27.5 shows the peptide bond formed between serine and alanine in the dipeptide serylalanine.

Peptide is the name given to a short polymer of amino acids. Peptides are classified by the number of amino acid units in the chain. A molecule containing 2 amino acids joined by an amide bond is called a **dipeptide.** Those containing 3 to 10 amino acids are called **tripeptides, tetrapeptides, pentapeptides,** and so on. Molecules containing more than 10 but fewer than 20 amino acids are called **oligopeptides.** Those containing several dozen or more amino acids are called **polypeptides. Proteins** are biological macromolecules of molecular weight 5000 or greater, consisting of one or more polypeptide chains. The distinctions in this terminology are not precise.

By convention, polypeptides are written from the left, beginning with the amino acid having the free $-NH_3^+$ group and proceeding to the right toward the amino

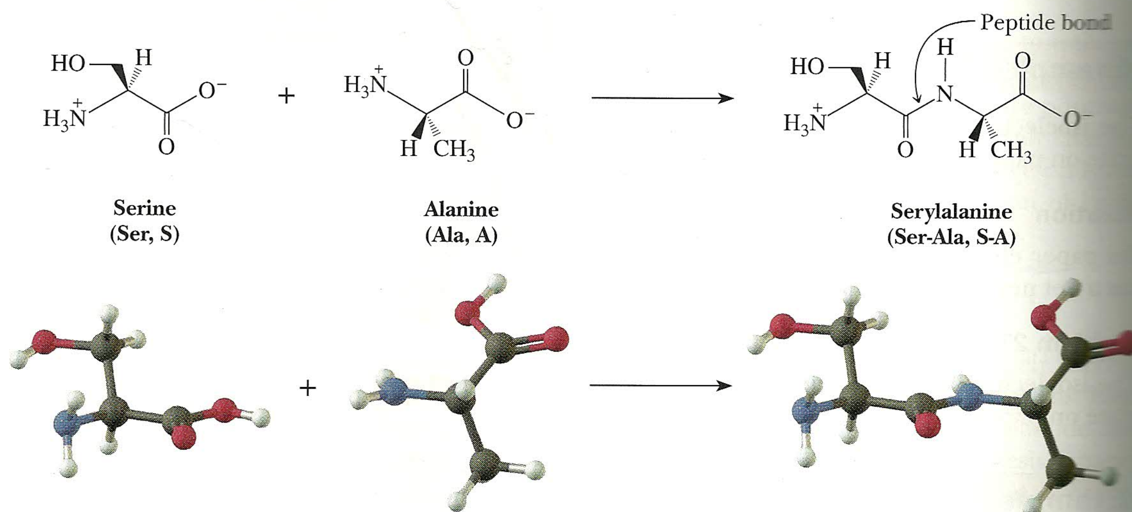
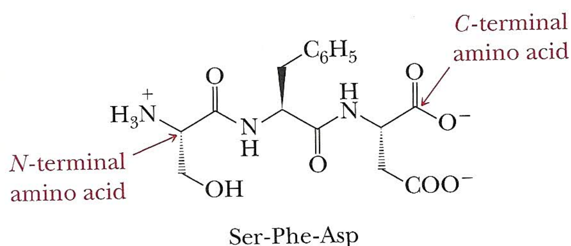


Figure 27.5
The peptide bond in serylalanine.

acid with the free —COO^- group. The amino acid with the free —NH_3^+ group is called the **N-terminal amino acid** and that with the free —COO^- group is called the **C-terminal amino acid.** Notice the repeating pattern in the peptide chain of N—α-carbon—carbonyl, etc.

N-Terminal amino acid The amino acid at the end of a polypeptide chain having the free —NH_2 group.

C-Terminal amino acid The amino acid at the end of a polypeptide chain having the free —COOH group.

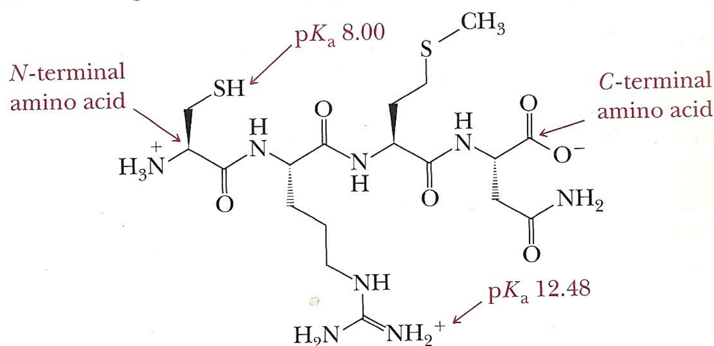


Example 27.4

Draw a structural formula for Cys-Arg-Met-Asn. Label the N-terminal amino acid and the C-terminal amino acid. What is the net charge on this tetrapeptide at pH 6.0?

Solution

The backbone of this tetrapeptide is a repeating sequence of nitrogen—α-carbon—carbonyl. The net charge on this tetrapeptide at pH 6.0 is +1.



Problem 27.4

Draw a structural formula for Lys-Phe-Ala. Label the N-terminal amino acid and the C-terminal amino acid. What is the net charge on this tripeptide at pH 6.0?

27.4 Primary Structure of Polypeptides and Proteins

The **primary (1°) structure** of a polypeptide or protein refers to the sequence of amino acids in its polypeptide chain. In this sense, primary structure is a complete description of all covalent bonding in a polypeptide or protein.

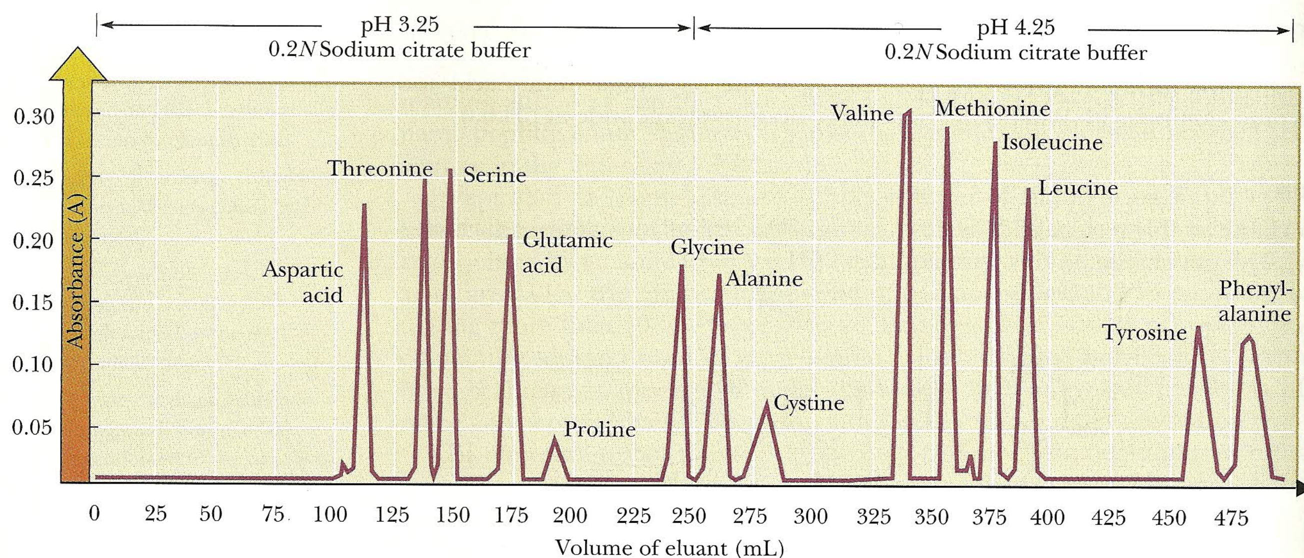
Primary structure of proteins The sequence of amino acids in the polypeptide chain, read from the N-terminal amino acid to the C-terminal amino acid.

In 1953, Frederick Sanger of Cambridge University, England, reported the primary structure of the two polypeptide chains of the hormone insulin. Not only was

this a remarkable achievement in analytical chemistry, but it also clearly established that the molecules of a given protein all have the same amino acid composition and the same amino acid sequence. Today, the amino acid sequences of over 20,000 different proteins are known.

A. Amino Acid Analysis

The first step for determining the primary structure of a polypeptide is hydrolysis and quantitative analysis of its amino acid composition. Recall from Section 18.4D that amide bonds are very resistant to hydrolysis. Typically, samples of protein are hydrolyzed in 6 *M* HCl in sealed glass vials at 110°C for 24 to 72 hours. This hydrolysis can be done in a microwave oven in a shorter time. After the polypeptide is hydrolyzed, the resulting mixture of amino acids is analyzed by ion-exchange chromatography.



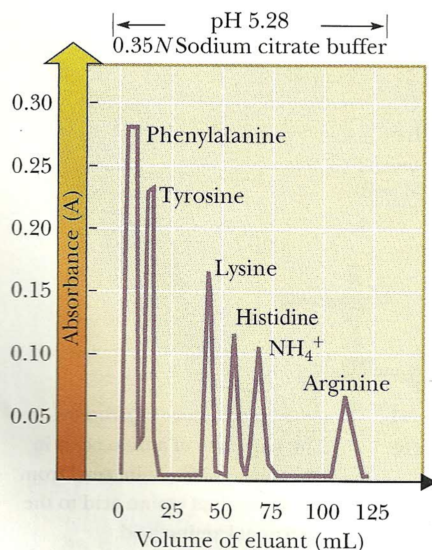


Figure 27.6
Analysis of a mixture of amino acids by ion-exchange chromatography using Amberlite IR-120, a sulfonated polystyrene resin. The resin contains phenyl-$SO_3^-Na^+$ groups. The amino acid mixture is applied to the column at low pH (3.25) under which conditions the acidic amino acids (Asp, Glu) are weakly bound to the resin, and the basic amino acids (Lys, His, Arg) are tightly bound. Sodium citrate buffers at two different concentrations, and three different values of pH are used to elute the amino acids from the column. Cysteine is determined as cystine, Cys-S-S-Cys, the disulfide of cysteine.

Amino acids are detected as they emerge from the column by reaction with ninhydrin (Section 27.2D) followed by absorption spectroscopy. Current procedures for hydrolysis of polypeptides and analysis of amino acid mixtures have been refined to the point where it is possible to obtain amino acid composition from as little as 50 nanomoles (50×10^{-9} mole) of polypeptide. Figure 27.6 shows the analysis of a polypeptide hydrolysate by ion-exchange chromatography. Note that during hydrolysis, the side-chain amide groups of asparagine and glutamine are hydrolyzed, and these amino acids are detected as aspartic acid and glutamic acid. For each glutamine or asparagine hydrolyzed, an equivalent amount of ammonium chloride is formed.

B. Sequence Analysis

After the amino acid composition of a polypeptide has been determined, the next step is to determine the order in which the amino acids are joined in the polypeptide chain. The most common sequencing strategy is to cleave the polypeptide at specific peptide bonds (using, for example, cyanogen bromide or certain proteolytic enzymes), determine the sequence of each fragment (using, for example, the Edman degradation), and then match overlapping fragments to arrive at the sequence of the polypeptide.

Cyanogen Bromide

Cyanogen bromide (BrCN) is specific for cleavage of peptide bonds formed by the carboxyl group of methionine (Figure 27.7). The products of this cleavage are a substituted γ-lactone (Section 18.1C) derived from the *N*-terminal portion of the polypeptide and a second fragment containing the *C*-terminal portion of the polypeptide.

A three-step mechanism can be written for this reaction. The strategy for cyanogen bromide cleavage depends on chemical manipulation of the leaving ability of the sulfur atom of methionine. Because CH_3S^- is the anion of a weak acid, it is a very poor leaving group, just as OH^- is a poor leaving group (Section 9.4F). Yet, just as the oxygen atom of an alcohol can be transformed into a better leaving group by converting it into an oxonium ion (by protonation), so too can the sulfur atom of methionine be transformed into a better leaving group by converting it into a sulfonium ion.

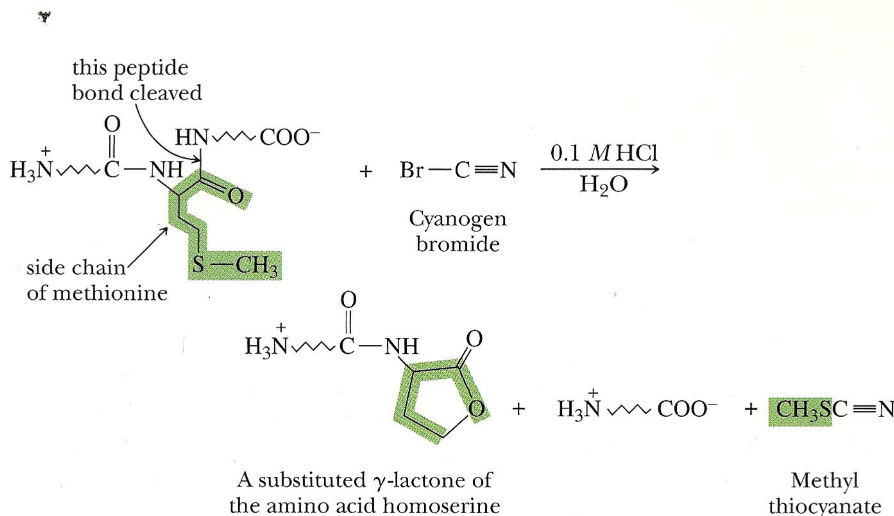
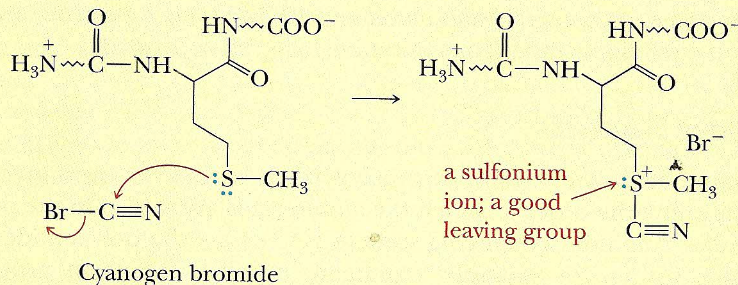


Figure 27.7 Cleavage by cyanogen bromide, BrCN, of a peptide bond formed by the carboxyl group of methionine.

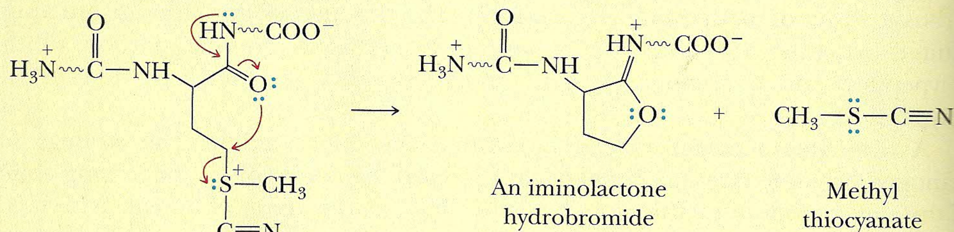
Mechanism *Cleavage of a Peptide Bond at Methionine by Cyanogen Bromide*

Step 1: Reaction is initiated by nucleophilic attack of the divalent sulfur atom of methionine on the carbon of cyanogen bromide displacing bromide ion. The product of this nucleophilic displacement is a sulfonium ion.



Cyanogen bromide

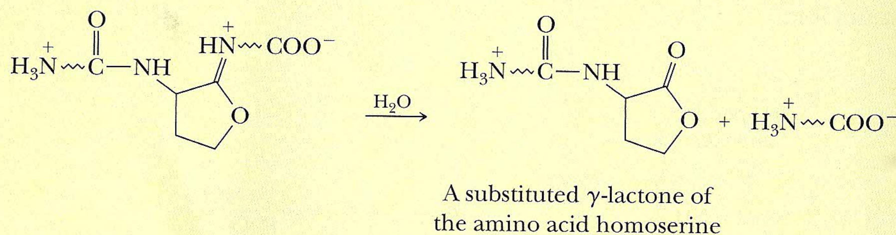
Step 2: An internal S_N2 reaction in which the oxygen of the methionine carbonyl group attacks the γ-carbon and displaces methyl thiocyanate gives a five-membered ring. Note that the oxygen of a carbonyl group is at best a weak nucleophile. This displacement is facilitated, however, because the sulfonium ion is a very good leaving group and because of the ease with which a five-membered ring is formed.



An iminolactone hydrobromide

Methyl thiocyanate

Step 3: Hydrolysis of the imino group gives a γ-lactone derived from the *N*-terminal end of the original polypeptide.



A substituted γ-lactone of the amino acid homoserine

Enzyme-Catalyzed Hydrolysis of Peptide Bonds

A group of proteolytic enzymes, among them trypsin and chymotrypsin, can be used to catalyze the hydrolysis of specific peptide bonds. Trypsin catalyzes the hydrolysis of peptide bonds formed by the carboxyl groups of arginine and lysine; chymotrypsin catalyzes the hydrolysis of peptide bonds formed by the carboxyl groups of phenylalanine, tyrosine, and tryptophan (Table 27.3).

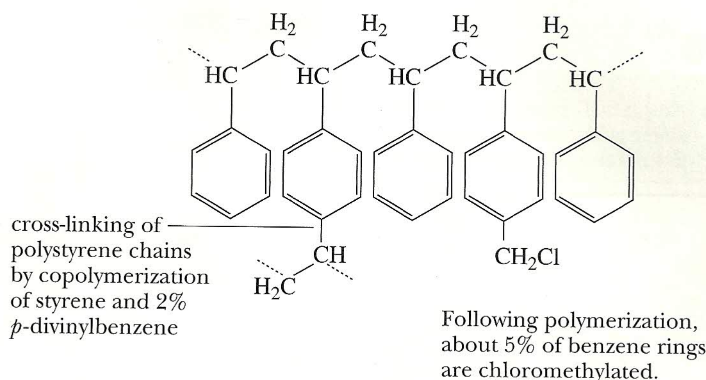
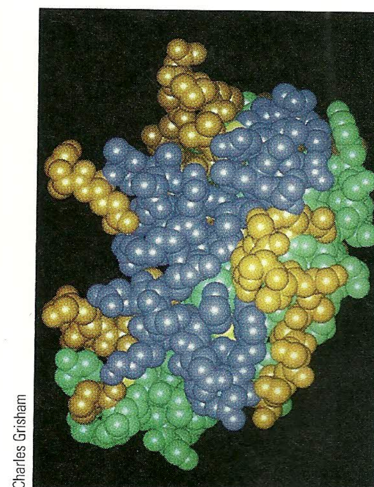


Figure 27.10 The support used for the Merrifield solid-phase synthesis is a chloromethylated polystyrene resin.

polypeptide is released from the polymer beads by cleavage of the benzyl ester. The steps in solid-phase synthesis of a polypeptide are summarized in Figure 27.11.

Thanks to automation, the synthesis of polypeptides is now a routine procedure in chemical research. It is common for researchers to order several peptides at a time for use in fields as diverse as medicine, biology, material science, and biomedical engineering.

A dramatic illustration of the power of the solid-phase method was the synthesis of the enzyme ribonuclease by Merrifield in 1969. The synthesis involved 369 chemical reactions and 11,931 operations, all of which were performed by an automated machine and without any intermediate isolation stages. Each of the 124 amino acids was added as an *N-tert*-butoxycarbonyl derivative and coupled using DCC. Cleavage from the resin and removal of all protective groups gave a mixture that was purified by ion-exchange chromatography. The specific activity of the synthetic enzyme was 13–24% of that of the natural enzyme. The fact that the specific activity of the synthetic enzyme was lower than that of the natural enzyme was probably attributable to the presence of polypeptide byproducts closely related to but not identical to the natural enzyme. Synthesizing ribonuclease (124 amino acids) requires forming 123 peptide bonds. If each peptide bond is formed in 99% yield, the yield of homogeneous polypeptide is $0.99^{123} = 29\%$. If each peptide bond is formed in 98% yield, the yield is 8%. Thus, even with yields as high as 99% in each peptide bond-forming step, a large portion of the synthetic polypeptides have one or more sequence defects. Many of these, nonetheless, may be fully or partially active.



A model of the protein ribonuclease A. The purple segments are regions of α-helix and the yellow segments are regions of β-pleated sheet, both of which are described in Section 27.6. Other colors represent loop regions.

27.6 Three-Dimensional Shapes of Polypeptides and Proteins

A. Geometry of a Peptide Bond

In the late 1930s, Linus Pauling began a series of studies to determine the geometry of a peptide bond. One of his first and most important discoveries was that a peptide bond itself is planar. As shown in Figure 27.12, the four atoms of a peptide bond and the two α-carbons joined to it all lie in the same plane.

Had you been asked in Chapter 1 to describe the geometry of a peptide bond, you probably would have predicted bond angles of 120° about the carbonyl carbon and 109.5° about the amide nitrogen. However, as fully discussed in Connections to Biological Chemistry: "The Unique Structure of Amide Bonds" in Chapter 18, both

Figure 27.11
Steps in the Merrifield solid-phase polypeptide synthesis.

1. Attach BOC-protected *C*-terminal amino acid to resin as benzyl ester. Ester protects its carboxyl group.

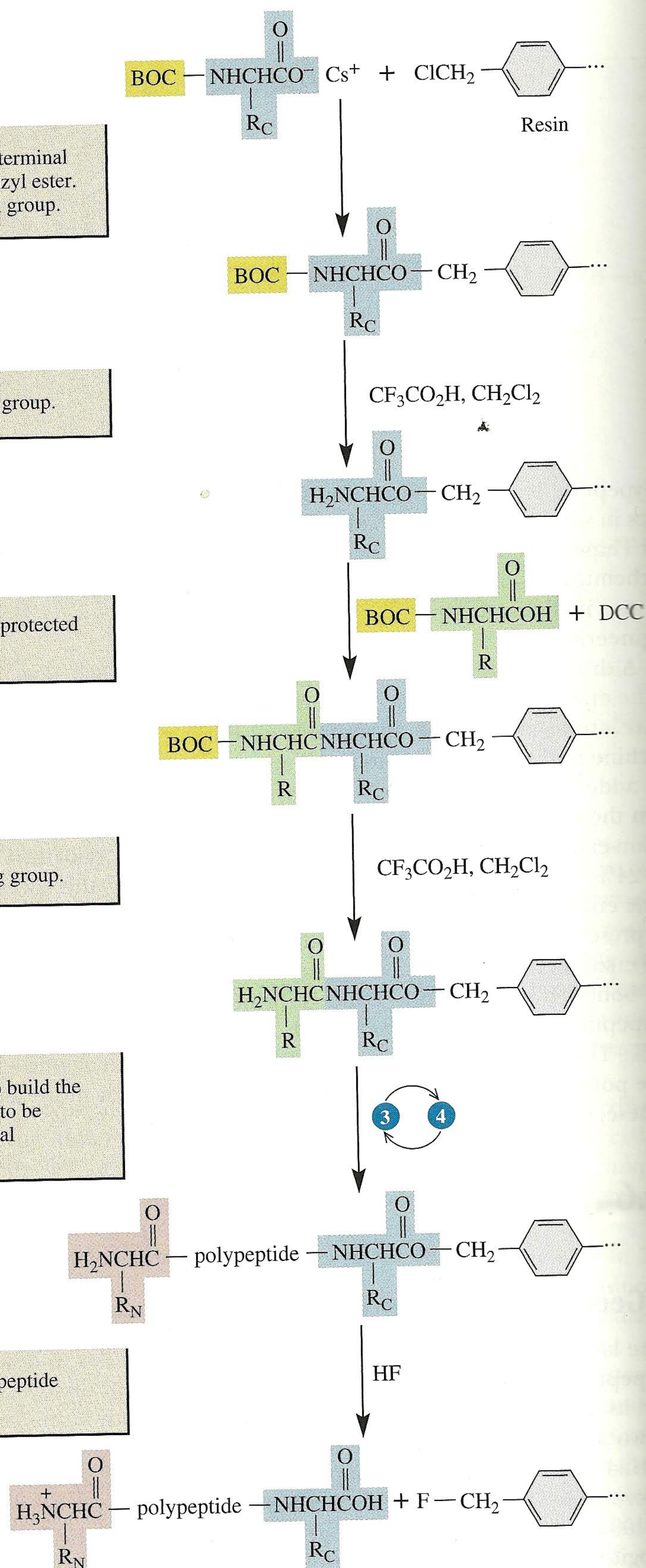
2. Remove BOC-protecting group.

3. Couple the second BOC-protected amino acid.

4. Remove BOC-protecting group.

Repeat 3 - 4 cycle to build the polypeptide chain. Last to be attached is the *N*-terminal amino acid.

5. Cleave completed polypeptide from resin.



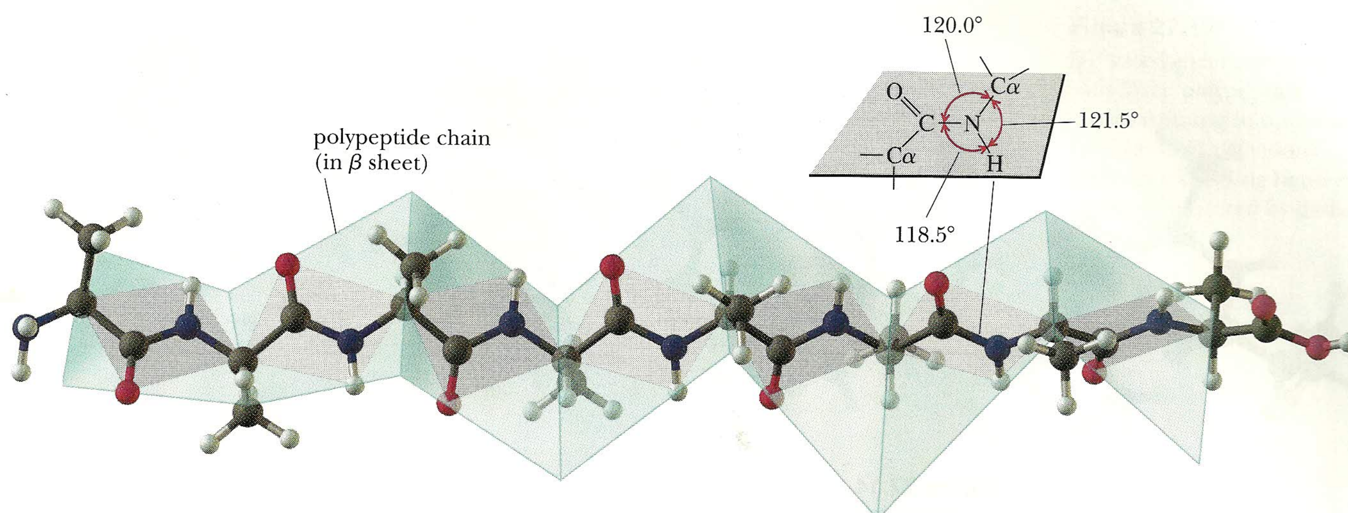
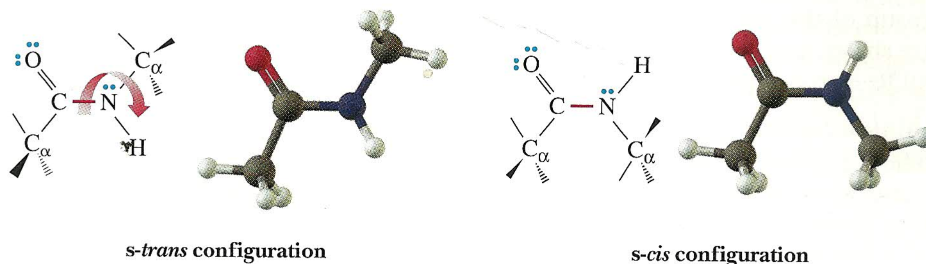


Figure 27.12
Planarity of a peptide bond. Bond angles about the carbonyl carbon and the amide nitrogen are approximately 120°.

atoms are actually planar with approximately 120° bond angles about each because of resonance of the nitrogen lone pair with the carbonyl. Two configurations are possible for the atoms of a planar peptide bond. In one, the two α-carbons are *cis* to each other; in the other, they are *trans* to each other. The *trans* configuration is more favorable because the α-carbons with the bulky groups bonded to them are farther from each other than they are in the *cis* configuration. Almost all peptide bonds in naturally occurring proteins studied to date have the *trans* configuration. Proline is found *cis* most of the time, and there are some well-known examples of other *cis* peptide bonds as well.



Secondary structure of proteins The ordered arrangements (conformations) of amino acids in localized regions of a polypeptide or protein.

B. Secondary Structure

Secondary (2°) structure refers to ordered arrangements (conformations) of amino acids in localized regions of a polypeptide or protein molecule. The first studies of polypeptide conformations were carried out by Linus Pauling and Robert Corey beginning in 1939. They assumed that in conformations of greatest stability, all atoms in a peptide bond lie in the same plane, and there is hydrogen bonding between the N—H of one peptide bond and the C=O of another, as shown in Figure 27.13.

On the basis of model building, Pauling proposed that two types of secondary structure should be particularly stable: the α-helix and the antiparallel β-pleated sheet. X-ray crystallography has validated this prediction completely.

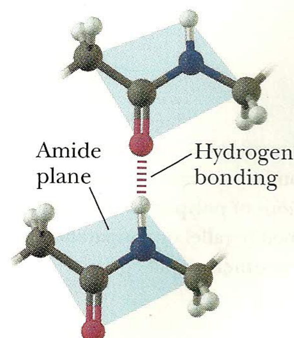


Figure 27.13
Hydrogen bonding between amide groups.

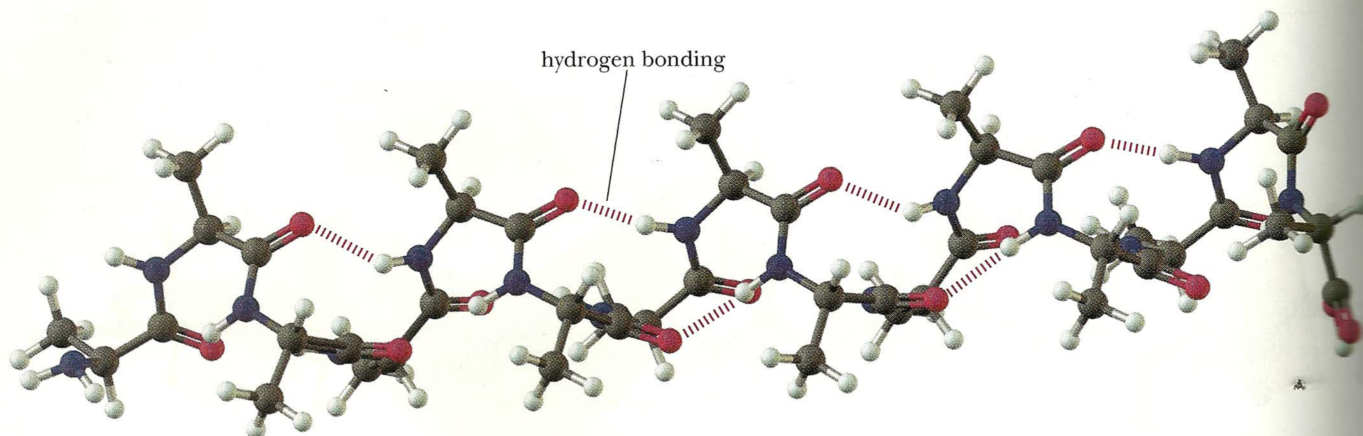


Figure 27.14
An α-helix. The peptide chain is repeating units of L-alanine.

α-Helix A type of secondary structure in which a section of polypeptide chain coils into a spiral, most commonly a right-handed spiral.

The α-Helix

In an **α-helix** pattern shown in Figure 27.14, a polypeptide chain is coiled in a spiral. As you study this section of α-helix, note the following.

1. The helix is coiled in a clockwise, or right-handed, manner. Right-handed means that if you turn the helix clockwise, it twists away from you. In this sense, a right-handed helix is analogous to the right-handed thread of a common wood or machine screw.
2. There are 3.6 amino acids per turn of the helix.
3. Each peptide bond is *trans* and planar.
4. The N—H group of each peptide bond points roughly downward, parallel to the axis of the helix, and the C=O of each peptide bond points roughly upward, also parallel to the axis of the helix.
5. The carbonyl group of each peptide bond is hydrogen-bonded to the N—H group of the peptide bond four amino acid units away from it. Hydrogen bonds are shown as dashed lines.
6. All R— groups point outward from the helix.

Almost immediately after Pauling proposed the α-helix conformation, other researchers proved the presence of α-helix conformations in keratin, the protein of hair and wool. It soon became obvious that the α-helix is one of the fundamental folding patterns of polypeptide chains.

β-Pleated sheet A type of secondary structure in which sections of polypeptide chains are aligned parallel or antiparallel to one another.

The β-Pleated Sheet

An antiparallel **β-pleated sheet** consists of extended polypeptide chains with neighboring chains running in opposite (antiparallel) directions. In a parallel β-pleated sheet, the polypeptide chains run in the same direction. Unlike the α-helix arrangement, N—H and C=O groups lie in the plane of the sheet and are roughly perpendicular to the long axis of the sheet. The C=O group of each peptide bond is hydrogen-bonded to the N—H group of a peptide bond of a neighboring chain (Figure 27.15).

As you study this section of β-pleated sheet, note the following.

1. The three polypeptide chains lie adjacent to each other and run in opposite (antiparallel) directions.
2. Each peptide bond is planar, and the α-carbons are *trans* to each other.

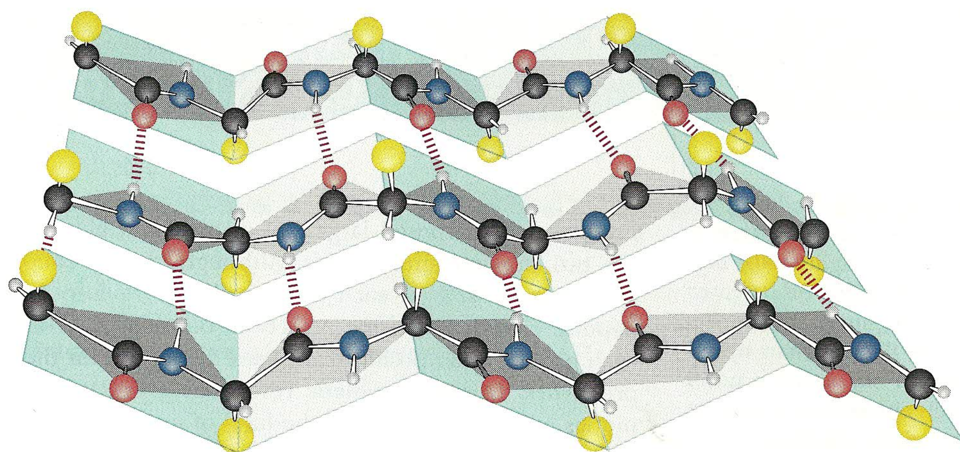


Figure 27.15
β-Pleated sheet conformation with three polypeptide chains running in opposite (antiparallel) directions. Hydrogen bonding between chains is indicated by dashed lines.

3. The C=O and N—H groups of peptide bonds from adjacent chains point at each other and are in the same plane so that hydrogen bonding is possible between adjacent polypeptide chains.
4. The R-groups on any one chain alternate, first above and then below the plane of the sheet, and so on.

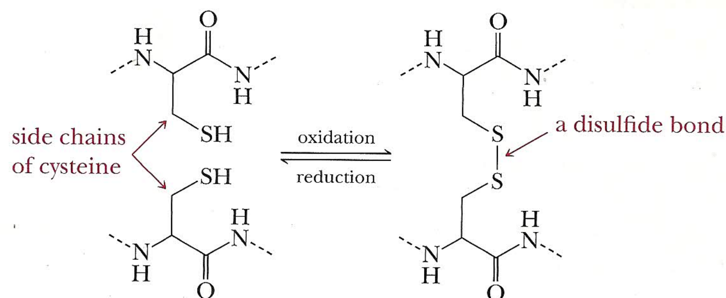
The *β*-pleated sheet conformation is stabilized by hydrogen bonding between N—H groups of one chain and C=O groups of an adjacent chain. By comparison, the *α*-helix is stabilized by hydrogen bonding between N—H and C=O groups within the same polypeptide chain.

C. Tertiary Structure

Tertiary (3°) structure refers to the overall folding pattern and arrangement in space of all atoms in a single polypeptide chain. No sharp dividing line exists between secondary and tertiary structures. Secondary structure refers to the spatial arrangement of amino acids close to one another on a polypeptide chain, whereas tertiary structure refers to the three-dimensional arrangement of all atoms of a polypeptide chain. Among the most important factors in maintaining 3° structure are disulfide bonds, hydrophobic interactions, hydrogen bonding, and salt linkages.

Tertiary structure of proteins The three-dimensional arrangement in space of all atoms in a single polypeptide chain.

Disulfide bonds (Section 10.9G) play an important role in maintaining tertiary structure. Disulfide bonds are formed between side chains of two cysteine units by oxidation of their thiol groups (—SH) to form a disulfide bond. Treatment of a disulfide bond with a reducing agent regenerates the thiol groups.



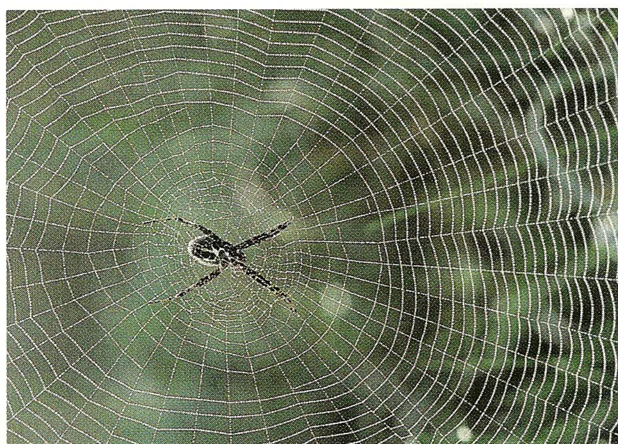
CHEMICAL CONNECTIONS

Spider Silk

Spider silk has some remarkable properties. Research is currently concentrated on the strong dragline silk that forms the spokes of a web of the Golden Orb Weaver (*Nephila clavipes*). This silk has three times the impact strength of Kevlar and is 30% more flexible than nylon. The commercial application of spider silk is not a novel concept. Eighteenth-century French entrepreneur Bon de Saint-Hilaire attempted to mass-produce silk in his high-density spider farms but failed because of cannibalism among his territorial arachnid workers. In contrast, native New Guineans continue to successfully collect and utilize spider silk for a wide range of applications including bags and fishing nets. Today, the only way to obtain large amounts of silk is to extract it from the abdomens of immobilized spiders, but scientific advances make the mass production and industrial application of spider silk increasingly possible.

Biologically produced dragline silk is a combination of two liquid proteins, Spidroin 1 and 2, which become oriented and solidify as they travel through a complex duct system in the spider's abdomen. These proteins are composed largely of alanine and glycine, the two smallest amino acids. Although glycine comprises almost 42% of each protein, the short, 5 to 10 peptide chains of alanine, which account for 25% of each protein's composition, are more important for the properties. Nuclear magnetic resonance (NMR) techniques have vastly improved the level of understanding of spider silk's structure, which was originally determined by x-ray crystallography. NMR data of spidroins containing deuterium-tagged alanine have shown that all alanines are configured into β-pleated sheets. Furthermore, the NMR data suggest that 40% of the alanine β-sheets are highly structured while the other 60% are less oriented, forming fingers that reach out from each individual strand. These fingers are believed to join the oriented alanine β-sheets and the glycine-rich, amorphous "background" sectors of the polypeptide.

Currently, genetically modified *Escherichia coli* is used to mass-produce Spidroin 1 and 2. However, DNA redundancy initially caused synthesis problems when the spider genes were transposed into the bacteria. The *E. coli* did not transcribe some of the codons in the same way that spider cells would, forcing scientists to modify the DNA. When the proteins could be synthesized, it was necessary to develop a system to mimic the natural production of spider silk while preventing the silk from contacting the air and subsequently hardening. After the two proteins are separated from the *E. coli,* they are drawn together into methanol through separate needles. Another approach is to dissolve the silk in formic acid or to add codons for hydrophilic amino acids, in this case histidine and arginine, to keep the artificial silk pliable. The industrial and practical applications of spider silk will not be fully known until it can be abiotically synthesized in large quantities.



Tom Bean/Stone/Getty

Golden orb weaver.

Based on a Chem 30H honors paper by Paul Celestre, UCLA.

Figure 27.16 shows the amino acid sequence of human insulin. This protein consists of two polypeptide chains: an A chain of 21 amino acids and a B chain of 30 amino acids. The A chain is bonded to the B chain by two interchain disulfide bonds. An intrachain disulfide bond also connects the cysteine units at positions 6 and 11 of the A chain.

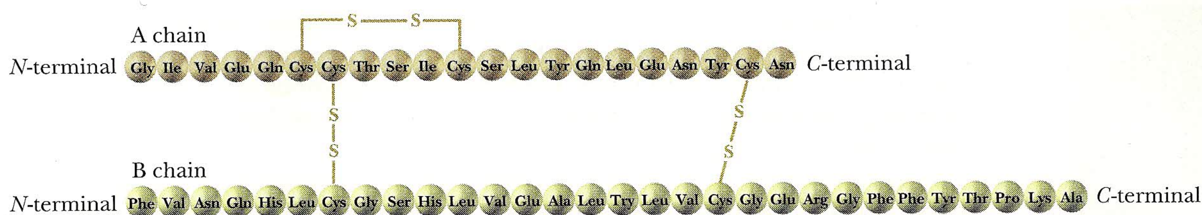


Figure 27.16
Human insulin. The A chain of 21 amino acids and B chain of 30 amino acids are connected by interchain disulfide bonds between A7 and B7 and between A20 and B19. In addition, a single intrachain disulfide bond occurs between A6 and A11.

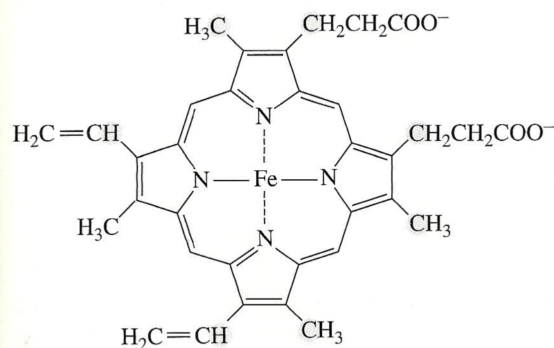
As an example of 2° and 3° structure, let us look at the three-dimensional structure of myoglobin—a protein found in skeletal muscle and particularly abundant in diving mammals, such as seals, whales, and porpoises. Myoglobin and its structural relative, hemoglobin, are the oxygen storage and transport molecules of vertebrates. Hemoglobin binds molecular oxygen in the lungs and transports it to myoglobin in muscles. Myoglobin stores molecular oxygen until it is required for metabolic oxidation.

Myoglobin consists of a single polypeptide chain of 153 amino acids. Myoglobin also contains a single heme unit. Heme consists of one Fe^{2+} ion coordinated in a square planar array with the four nitrogen atoms of a molecule of porphyrin (Figure 27.17).

Determination of the three-dimensional structure of myoglobin represented a milestone in the study of molecular architecture. For their contribution to this research, John C. Kendrew and Max F. Perutz, both of Britain, shared the 1962 Nobel Prize for chemistry. The secondary and tertiary structures of myoglobin are shown in Figure 27.18. The single polypeptide chain is folded into a complex, almost boxlike shape.

Following are important structural features of the three-dimensional shape of myoglobin.

1. The backbone consists of eight relatively straight sections of α-helix, each separated by a bend in the polypeptide chain. The longest section of α-helix has 24 amino acids, the shortest has 7. Some 75% of the amino acids are found in these eight regions of α-helix.



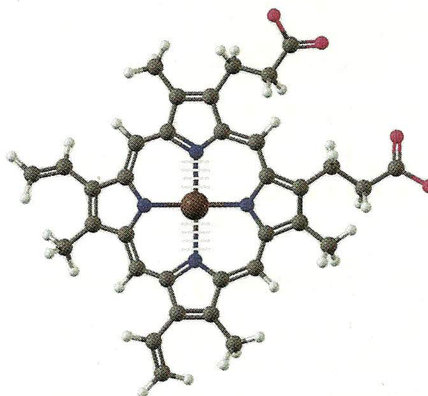
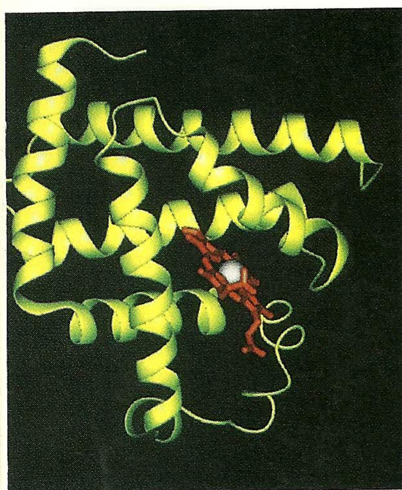


Figure 27.17
The structure of heme, found in myoglobin and hemoglobin.



Brent Iverson, University of Texas

Figure 27.18
Ribbon model of myoglobin. The polypeptide chain is shown in yellow, the heme ligand in red, and the Fe atom as a white sphere.

2. Hydrophobic side chains of phenylalanine, alanine, valine, leucine, isoleucine, and methionine are clustered in the interior of the molecule where they are shielded from contact with water. **Hydrophobic interactions** are a major factor in directing the folding of the polypeptide chain of myoglobin into this compact, three-dimensional shape.
3. The outer surface of myoglobin is coated with hydrophilic side chains, such as those of lysine, arginine, serine, glutamic acid, histidine, and glutamine, which interact with the aqueous environment by **hydrogen bonding.** The only polar side chains that point to the interior of the myoglobin molecule are those of two histidine units, which point inward toward the heme group.
4. Oppositely charged amino acid side chains close to each other in the three-dimensional structure interact by electrostatic attractions called **salt linkages.** An example of a salt linkage is the attraction of the side chains of lysine ($—NH_3^+$) and glutamic acid ($—COO^-$).

The tertiary structures of hundreds of proteins have also been determined. It is clear that proteins contain α-helix and β-pleated sheet structures, but that wide variations exist in the relative amounts of each. Lysozyme, with 129 amino acids in a single polypeptide chain, has only 25% of its amino acids in α-helix regions. Cytochrome, with 104 amino acids in a single polypeptide chain, has no α-helix structure but does contain several regions of β-pleated sheet. Yet, whatever the proportions of α-helix, β-pleated sheet, or other periodic structure, most nonpolar side chains of water-soluble proteins are directed toward the interior of the molecule, whereas polar side chains are on the surface of the molecule and in contact with the aqueous environment. Note that this arrangement of polar and nonpolar groups in water-soluble proteins very much resembles the arrangement of polar and nonpolar groups of soap molecules in micelles (Figure 26.3). It also resembles the arrangement of phospholipids in lipid bilayers (Figure 26.13).

Example 27.7

With which of the following amino acid side chains can the side chain of threonine form hydrogen bonds?

(a) Valine **(b)** Asparagine **(c)** Phenylalanine
(d) Histidine **(e)** Tyrosine **(f)** Alanine

Solution

The side chain of threonine contains a hydroxyl group that can participate in hydrogen bonding in two ways: Its oxygen has a partial negative charge and can function as a hydrogen bond acceptor, and its hydrogen has a partial positive charge and can function as a hydrogen bond donor. Therefore, the side chain of threonine can form hydrogen bonds with the side chains of tyrosine, asparagine, and histidine.

Problem 27.7

At pH 7.4, with what amino acid side chains can the side chain of lysine form salt linkages?

D. Quaternary Structure

Quaternary structure The arrangement of polypeptide monomers into a noncovalently bonded aggregate.

Most proteins of molecular weight greater than 50,000 consist of two or more noncovalently linked polypeptide chains. The arrangement of protein monomers into an aggregation is known as **quaternary (4°) structure.** A good example is hemoglobin, a

protein that consists of four separate polypeptide chains: two α-chains of 141 amino acids each and two β-chains of 146 amino acids each. The quaternary structure of hemoglobin is shown in Figure 27.19.

A major factor stabilizing the aggregation of protein subunits is the **hydrophobic effect.** When separate polypeptide chains fold into compact three-dimensional shapes to expose polar side chains to the aqueous environment and shield nonpolar side chains from water, hydrophobic "patches" still appear on the surface, in contact with water. These patches can be shielded from water if two or more monomers assemble so that their hydrophobic patches are in contact. The numbers of subunits of several proteins of known quaternary structure are shown in Table 27.4. Other important factors include correctly located complementary hydrogen bonding and charged sites on different subunits. The formation of aggregates of well-defined structure based on specific structural units on the subunits is being explored in the new field of molecular recognition.

Hydrophobic effect The tendency of nonpolar groups to cluster to shield themselves from contact with an aqueous environment.

Table 27.4 Quaternary Structure of Selected Proteins

Protein	Number of Subunits
Alcohol dehydrogenase	2
Aldolase	4
Hemoglobin	4
Lactate dehydrogenase	4
Insulin	6
Glutamine synthetase	12
Tobacco mosaic virus protein disc	17

Brent Iverson, University of Texas

Figure 27.19
Ribbon model of hemoglobin. The α-chains are shown in purple, the β-chains in yellow, the heme ligands in red, and the Fe atoms as white spheres.

Summary

Amino acids are compounds that contain both an amino group and a carboxyl group (Section 27.1A). A **zwitterion** is an internal salt of an amino acid. With the exception of glycine, all protein-derived amino acids are chiral (Section 27.1B). In the D,L convention, all are L-amino acids. In the *R,S* convention, 18 are (*S*)-amino acids. Although cysteine has the same absolute configuration, it is an (*R*)-amino acid because of the manner in which priorities are assigned about the tetrahedral chiral center. Isoleucine and threonine contain a second chiral center. The 20 protein-derived amino acids are commonly divided into four categories (Section 27.1C): nine with nonpolar side chains, four with polar but un-ionized side chains, four with acidic side chains, and three with basic side chains.

The **isoelectric point, pI,** of an amino acid, polypeptide, or protein is the pH at which it has no net charge (Section 27.2C). **Electrophoresis** is the process of separating compounds on the basis of their electric charge (Section 27.2D). Compounds having a high charge density move more rapidly than those with a lower charge density. Any amino acid or protein in a solution with a pH that equals the pI of the compound remains at the origin.

A **peptide bond** is the special name given to the amide bond formed between α-amino acids (Section 27.3). A **polypeptide** is a biological macromolecule containing many amino acids, each joined to the next by a peptide bond. By convention, the sequence of amino acids in a polypeptide is written beginning with the ***N*-terminal amino acid** toward the ***C*-terminal amino acid. Primary (1°) structure** of a polypeptide is the sequence of amino acids in the polypeptide chain (Section 27.4).

In solid-phase synthesis (Section 27.5F), or polymer-supported synthesis of polypeptides, the *C*-terminal amino acid is joined to a chloromethylated polystyrene resin as a benzyl ester. The polypeptide chain is then extended one amino acid at a time from the *N*-terminal end. When synthesis is completed,