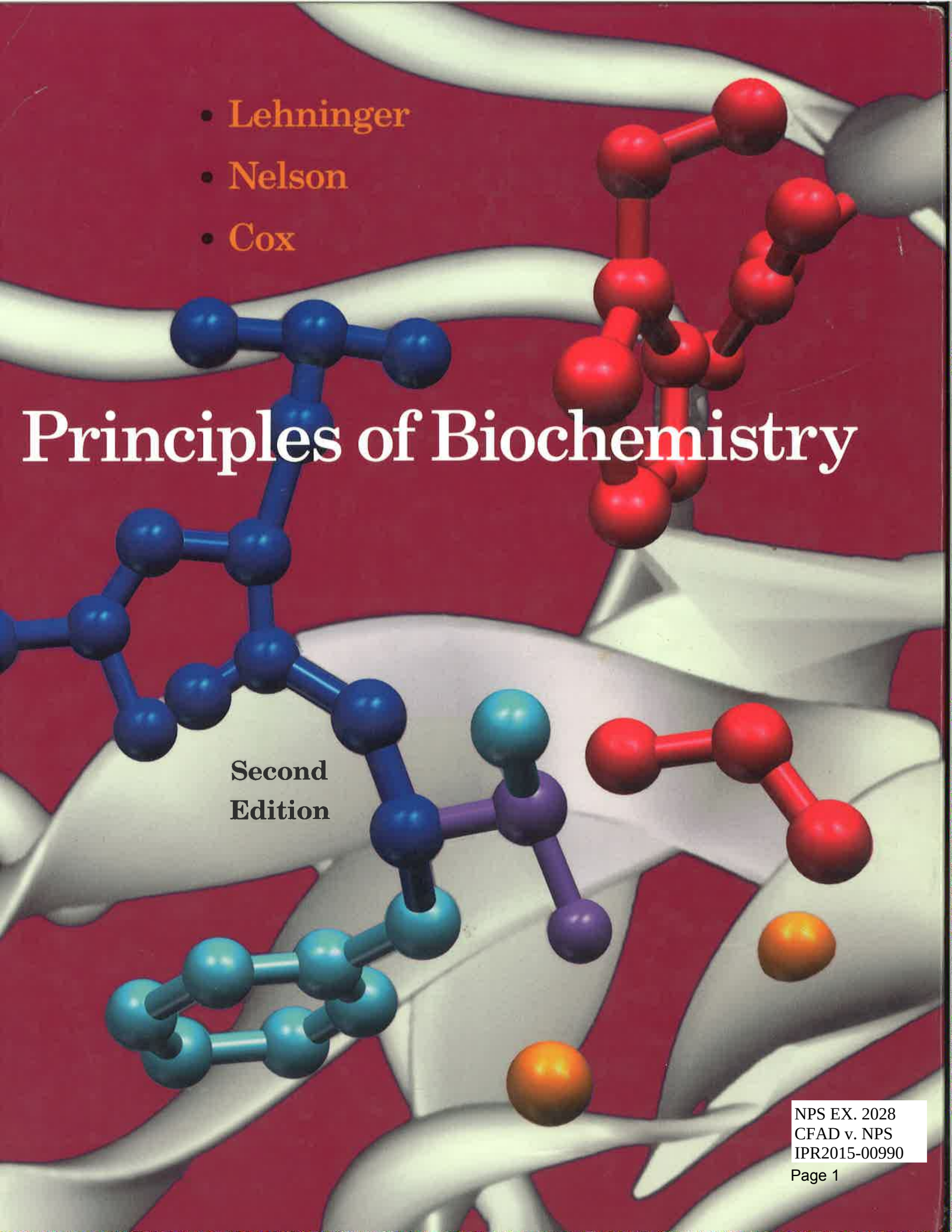


**Principles of Biochemistry**

Second Edition

Albert L. Lehninger, David L. Nelson, and Michael M. Cox



Printed in the United States of America

Library of Congress Catalog Card No. 91-67492

ISBN: 0-87901-500-4

Printing: 5 4 3 2 Year: 97 96 95 94 93

Development Editor: Valerie Neal

Design: Malcolm Grear Designers

Art Director: George Touloumes

Project Editor: Elizabeth Geller

Production Supervisor: Sarah Segal

Layout: Patricia Lawson

Picture Editor: Stuart Kenter

Illustration Design: Susan Tilberry

Illustrators: Susan Tilberry, Alan Landau, and Joan Waites

Computer Art: Laura Pardi Duprey and York Graphic Services

Composition: York Graphic Services

Printing and binding: R.R. Donnelley and Sons

Cover: The active site of the proteolytic enzyme chymotrypsin, showing the substrate (blue and purple) and the amino acid residues (red and orange) critical to catalysis. Determination of the detailed reaction mechanism of this enzyme (described on pp. 223–226) helped to establish the general principles of enzyme action.

Frontispiece: A view of tobacco ribulose-1,5-bisphosphate carboxylase (rubisco). This enzyme is central to photosynthetic carbon dioxide fixation; it is the most abundant enzyme in the biosphere. Different subunits are shown in blues and grays. Important active site residues are shown in red. Sulfates bound at the active site (an artifact of the crystallization procedure) are shown in yellow.

Cover, frontispiece, and part opening images produced by Alisa Zapp (see *Molecular Modeling* credits, p. IC-4) and enhanced by Academy Arts.

Illustration credits begin on p. IC-1 and constitute a continuation of the copyright page.

**Worth Publishers**

33 Irving Place

New York, NY 10003

CHAPTER

# Amino Acids and Peptides

Proteins are the most abundant macromolecules in living cells, occurring in all cells and all parts of cells. Proteins also occur in great variety; thousands of different kinds may be found in a single cell. Moreover, proteins exhibit great diversity in their biological function. Their central role is made evident by the fact that proteins are the most important final products of the information pathways discussed in Part IV of this book. In a sense, they are the molecular instruments through which genetic information is expressed. It is appropriate to begin the study of biological macromolecules with the proteins, whose name derives from the Greek *prōtos,* meaning "first" or "foremost."

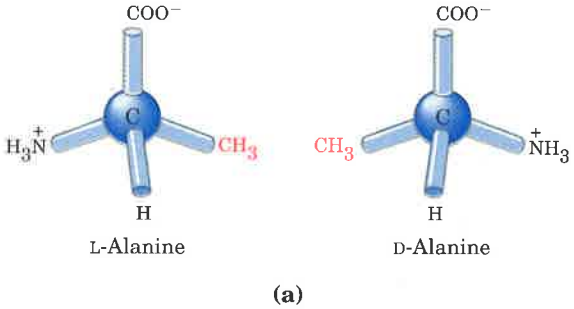
Relatively simple monomeric subunits provide the key to the structure of the thousands of different proteins. All proteins, whether from the most ancient lines of bacteria or from the most complex forms of life, are constructed from the same ubiquitous set of 20 amino acids, covalently linked in characteristic linear sequences. Because each of these amino acids has a distinctive side chain that determines its chemical properties, this group of 20 precursor molecules may be regarded as the alphabet in which the language of protein structure is written.

Proteins are chains of amino acids, each joined to its neighbor by a specific type of covalent bond. What is most remarkable is that cells can produce proteins that have strikingly different properties and activities by joining the same 20 amino acids in many different combinations and sequences. From these building blocks different organisms can make such widely diverse products as enzymes, hormones, antibodies, the lens protein of the eye, feathers, spider webs, rhinoceros horns (Fig. 5–1), milk proteins, antibiotics, mushroom poisons, and a myriad of other substances having distinct biological activities.

Protein structure and function is the topic for the next four chapters. In this chapter we begin with a description of amino acids and the covalent bonds that link them together in peptides and proteins.

**Figure 5–1** The protein keratin is formed by all vertebrates. It is the chief structural component of hair, scales, horn, wool, nails, and feathers. The black rhinoceros is nearing extinction in the wild because of the myths prevalent in some parts of the world that a powder derived from its horn has aphrodisiac properties. In reality, the chemical properties are no different from those of powdered bovine hooves or human fingernails.


**Figure 5–2** General structure of the amino acids found in proteins. With the exception of the nature of the R group, this structure is common to all the α-amino acids. (Proline, because it is an imino acid, is an exceptional component of proteins.) The α carbon is shown in blue. R (in red) represents the R group or side chain, which is different in each amino acid. In all amino acids except glycine (shown for comparison) the α-carbon atom has four different substituent groups.




**(a)**


**(b)**


**(c)**

## Amino Acids

Proteins can be reduced to their constituent amino acids by a variety methods, and the earliest studies of proteins naturally focused on t free amino acids derived from them. The first amino acid to be disco ered in proteins was asparagine, in 1806. The last of the 20 to be foun threonine, was not identified until 1938. All the amino acids have tri ial or common names, in some cases derived from the source fro which they were first isolated. Asparagine was first found in aspar gus, as one might guess; glutamate was found in wheat gluten; tyr sine was first isolated from cheese (thus its name is derived from t Greek *tyros,* "cheese"); and glycine (Greek *glykos,* "sweet") was named because of its sweet taste.

### Amino Acids Have Common Structural Features

All of the 20 amino acids found in proteins have a carboxyl group a an amino group bonded to the same carbon atom (the α carbon) (Fi 5–2). They differ from each other in their side chains, or R group which vary in structure, size, and electric charge, and influence t solubility of amino acids in water. When the R group contains ad tional carbons in a chain, they are designated β, γ, δ, ϵ, etc., proceedi out from the α carbon. The 20 amino acids of proteins are often referr to as the standard, primary, or normal amino acids, to distingui them from amino acids within proteins that are modified after t proteins are synthesized, and from many other kinds of amino aci present in living organisms but not in proteins. The standard ami acids have been assigned three-letter abbreviations and one-lett symbols (Table 5–1), which are used as shorthand to indicate the co position and sequence of amino acids in proteins.

We note in Figure 5–2 that for all the standard amino acids exce one (glycine) the α carbon is asymmetric, bonded to four different su stituent groups: a carboxyl group, an amino group, an R group, and hydrogen atom. The α-carbon atom is thus a **chiral center** (see F 3–9). Because of the tetrahedral arrangement of the bonding orbita around the α-carbon atom of amino acids, the four different substit ent groups can occupy two different arrangements in space, which a nonsuperimposable mirror images of each other (Fig. 5–3). These t forms are called **enantiomers** or **stereoisomers** (see Fig. 3–9). molecules with a chiral center are also **optically active**—i.e., th can rotate plane-polarized light, with the direction of the rotation d fering for different stereoisomers.

**Figure 5–3** **(a)** The two stereoisomers of alanine. L- and D-alanine are nonsuperimposable mirror images of each other. **(b, c)** Two different conventions for showing the configurations in space of stereoisomers. In perspective formulas **(b)** the wedge-shaped bonds project out of the plane of the paper, the dashed bonds behind it. In projection formulas **(c)** the horizontal bonds are assumed to project out of the plane of the paper, the vertical bonds behind. However, projection formulas are often used casually without reference to stereochemical configuration.

**Table 5–1 Properties and conventions associated with the standard amino acids**

| Amino acid | Abbreviated names | | $M_r$ | p$K_1$ (—COOH) | p$K_2$ (—$NH_3^+$) | p$K_R$ (R group) | pI | Hydropathy index* | Occurrence in Proteins (%)† |
|---|---|---|---|---|---|---|---|---|---|
| *Nonpolar, aliphatic R groups* | | | | | | | | | |
| Glycine | Gly | G | 75 | 2.34 | 9.60 | | 5.97 | −0.4 | 7.5 |
| Alanine | Ala | A | 89 | 2.34 | 9.69 | | 6.01 | 1.8 | 9.0 |
| Valine | Val | V | 117 | 2.32 | 9.62 | | 5.97 | 4.2 | 6.9 |
| Leucine | Leu | L | 131 | 2.36 | 9.60 | | 5.98 | 3.8 | 7.5 |
| Isoleucine | Ile | I | 131 | 2.36 | 9.68 | | 6.02 | 4.5 | 4.6 |
| Proline | Pro | P | 115 | 1.99 | 10.96 | | 6.48 | −1.6 | 4.6 |
| *Aromatic R groups* | | | | | | | | | |
| Phenylalanine | Phe | F | 165 | 1.83 | 9.13 | | 5.48 | 2.8 | 3.5 |
| Tyrosine | Tyr | Y | 181 | 2.20 | 9.11 | 10.07 | 5.66 | −1.3 | 3.5 |
| Tryptophan | Trp | W | 204 | 2.38 | 9.39 | | 5.89 | −0.9 | 1.1 |
| *Polar, uncharged R groups* | | | | | | | | | |
| Serine | Ser | S | 105 | 2.21 | 9.15 | 13.60 | 5.68 | −0.8 | 7.1 |
| Threonine | Thr | T | 119 | 2.11 | 9.62 | 13.60 | 5.87 | −0.7 | 6.0 |
| Cysteine | Cys | C | 121 | 1.96 | 8.18 | 10.28 | 5.07 | 2.5 | 2.8 |
| Methionine | Met | M | 149 | 2.28 | 9.21 | | 5.74 | 1.9 | 1.7 |
| Asparagine | Asn | N | 132 | 2.02 | 8.80 | | 5.41 | −3.5 | 4.4 |
| Glutamine | Gln | Q | 146 | 2.17 | 9.13 | | 5.65 | −3.5 | 3.9 |
| *Negatively charged R groups* | | | | | | | | | |
| Aspartate | Asp | D | 133 | 1.88 | 9.60 | 3.65 | 2.77 | −3.5 | 5.5 |
| Glutamate | Glu | E | 147 | 2.19 | 9.67 | 4.25 | 3.22 | −3.5 | 6.2 |
| *Positively charged R groups* | | | | | | | | | |
| Lysine | Lys | K | 146 | 2.18 | 8.95 | 10.53 | 9.74 | −3.9 | 7.0 |
| Arginine | Arg | R | 174 | 2.17 | 9.04 | 12.48 | 10.76 | −4.5 | 4.7 |
| Histidine | His | H | 155 | 1.82 | 9.17 | 6.00 | 7.59 | −3.2 | 2.1 |

* A scale combining hydrophobicity and hydrophilicity; can be used to predict which amino acids will be found in an aqueous environment (− values) and which will be found in a hydrophobic environment (+ values). See Box 10-2. From Kyte, J. & Doolittle, R.F. (1982) *J. Mol. Biol.* **157,** 105–132.

† Average occurrence in over 200 proteins. From Klapper, M.H. (1977) *Biochem. Biophys. Res. Commun.* **78,** 1018–1024.

The classification and naming of stereoisomers is based on the **absolute configuration** of the four substituents of the asymmetric carbon atom. For this purpose a reference compound has been chosen, to which all other optically active compounds are compared. This reference compound is the 3-carbon sugar glyceraldehyde (Fig. 5–4), the smallest sugar to have an asymmetric carbon atom. The naming of configurations of both simple sugars and amino acids is based on the absolute configuration of glyceraldehyde, as established by x-ray diffraction analysis. The stereoisomers of all chiral compounds having a configuration related to that of L-glyceraldehyde are designated L (for levorotatory, derived from *levo*, meaning "left"), and the stereoisomers related to D-glyceraldehyde are designated D (for dextrorotatory, derived from *dextro*, meaning "right"). The symbols L and D thus refer to the absolute configuration of the four substituents around the chiral carbon.

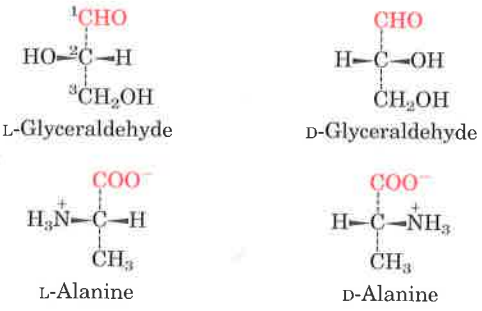



**Figure 5–4** Steric relationship of the stereoisomers of alanine to the absolute configuration of L- and D-glyceraldehyde. In these perspective formulas, the carbons are lined up vertically, with the chiral atom in the center. The carbons in these molecules are numbered beginning with the aldehyde or carboxyl carbons on the end, or 1 to 3 from top to bottom as shown. When presented in this way, the R group of the amino acid (in this case the methyl group of alanine) is always below the $\alpha$ carbon. L-Amino acids are those with the $\alpha$-amino group on the left, and D-amino acids have the $\alpha$-amino group on the right.

## Proteins Contain L-Amino Acids

Nearly all biological compounds with a chiral center occur naturally in only one stereoisomeric form, either D or L. The amino acids in protein molecules are the L stereoisomers. D-Amino acids have been found only in small peptides of bacterial cell walls and in some peptide antibiotics (see Fig. 5–19).

It is remarkable that the amino acids of proteins are all L stereoisomers. As we noted in Chapter 3, when chiral compounds are formed by ordinary chemical reactions, a racemic mixture of D and L isomers results. Whereas the L and D forms of chiral molecules are difficult for a chemist to distinguish and isolate, they are as different as night and day to a living system. The ability of cells to specifically synthesize the L isomer of amino acids reflects one of many extraordinary properties of enzymes (Chapter 8). The stereospecificity of the reactions catalyzed by some enzymes is made possible by the asymmetry of their active sites. The characteristic three-dimensional structures of proteins (Chapter 7), which dictate their diverse biological activities, require that *all* their constituent amino acids be of one stereochemical series.

## Amino Acids Are Ionized in Aqueous Solutions

Amino acids in aqueous solution are ionized and can act as acids or bases. Knowledge of the acid–base properties of amino acids is extremely important in understanding the physical and biological properties of proteins. Moreover, the technology of separating, identifying, and quantifying the different amino acids, which are necessary steps in determining the amino acid composition and sequence of protein molecules, is based largely on their characteristic acid–base behavior.

Those $\alpha$-amino acids having a single amino group and a single carboxyl group crystallize from neutral aqueous solutions as fully ionized species known as **zwitterions** (German for "hybrid ions"), each having both a positive and a negative charge (Fig. 5–5). These ions are electrically neutral and remain stationary in an electric field. The dipolar nature of amino acids was first suggested by the observation that crystalline amino acids have melting points much higher than those of other organic molecules of similar size. The crystal lattice of amino acids is held together by strong electrostatic forces between positively and negatively charged functional groups of neighboring molecules, resembling the stable ionic crystal lattice of NaCl (see Fig. 4–6).


**Figure 5–5** Nonionic and zwitterionic forms of amino acids. Note the separation of the + and − charges in the zwitterion, which makes it an electric dipole. The nonionic form does not occur in significant amounts in aqueous solutions. The zwitterion predominates at neutral pH.

## Amino Acids Can Be Classified by R Group

An understanding of the chemical properties of the standard amino acids is central to an understanding of much of biochemistry. The topic can be simplified by grouping the amino acids into classes based on the properties of their R groups (Table 5–1), in particular, their **polarity** or tendency to interact with water at biological pH (near pH 7.0). The polarity of the R groups varies widely, from totally nonpolar or hydrophobic (water-insouble) to highly polar or hydrophilic (water-soluble).

The structures of the 20 standard amino acids are shown in Figure 5–6, and many of their properties are listed in Table 5–1. There are five main classes of amino acids, those whose R groups are: nonpolar and aliphatic; aromatic (generally nonpolar); polar but uncharged; negatively charged; and positively charged. Within each class there are gradations of polarity, size, and shape of the R groups.

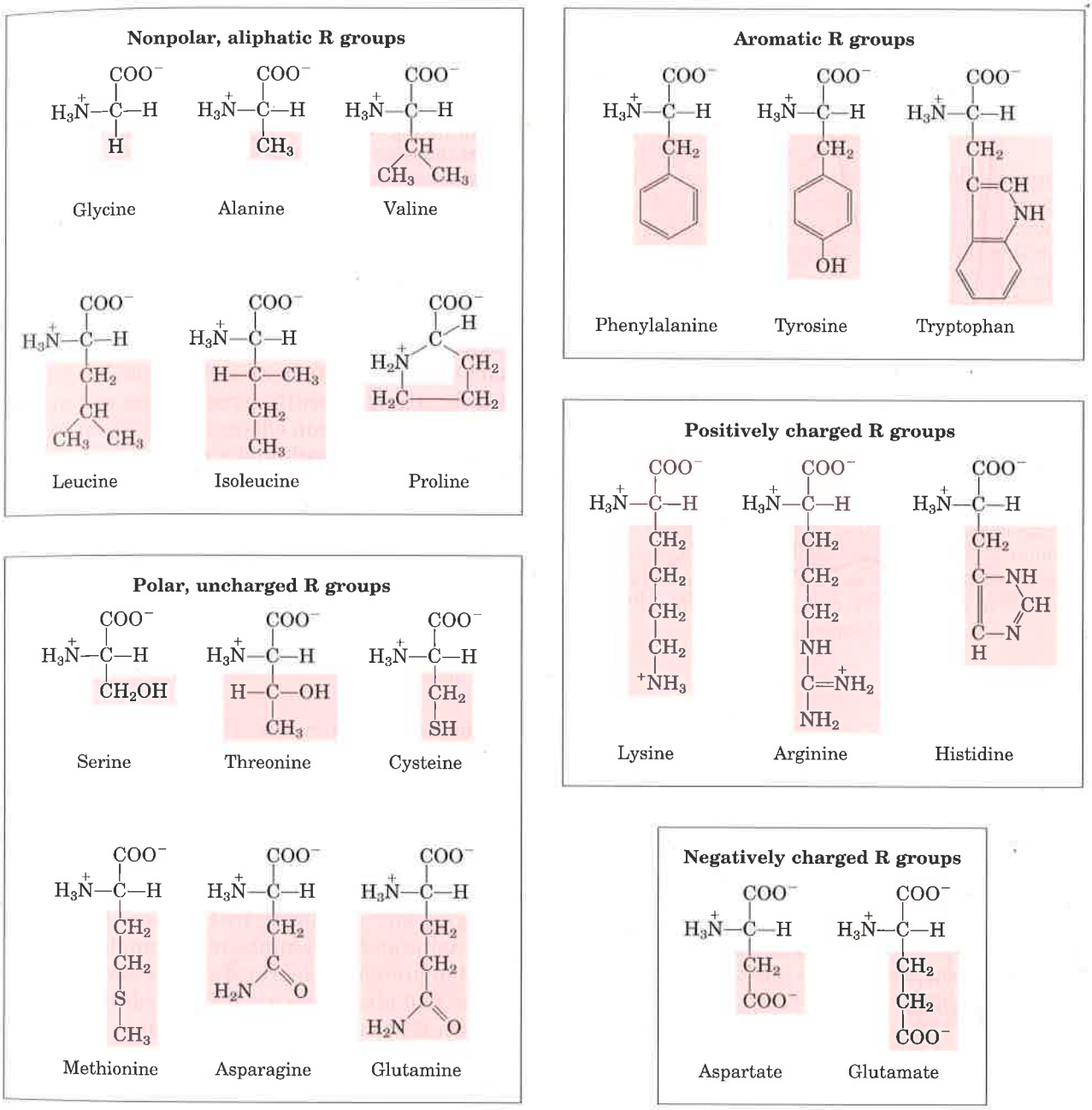



**Figure 5–6** The 20 standard amino acids of proteins. They are shown with their amino and carboxyl groups ionized, as they would occur at pH 7.0. The portions in black are those common to all the amino acids; the portions shaded in red are the R groups.

***Nonpolar, Aliphatic R Groups*** The hydrocarbon R groups in this class of amino acids are nonpolar and hydrophobic (Fig. 5–6). The bulky side chains of **alanine, valine, leucine,** and **isoleucine,** with their distinctive shapes, are important in promoting hydrophobic interactions within protein structures. **Glycine** has the simplest amino acid structure. Where it is present in a protein, the minimal steric hindrance of the glycine side chain allows much more structural flexibility than the other amino acids. **Proline** represents the opposite structural extreme. The secondary amino (imino) group is held in a rigid conformation that reduces the structural flexibility of the protein at that point.

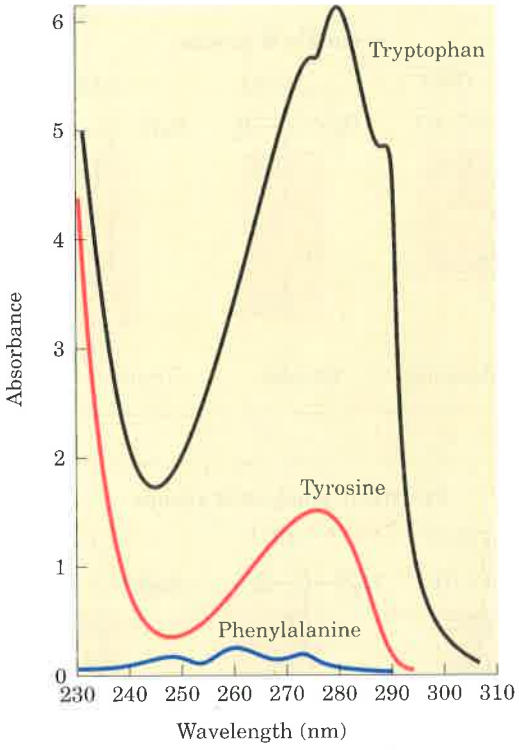


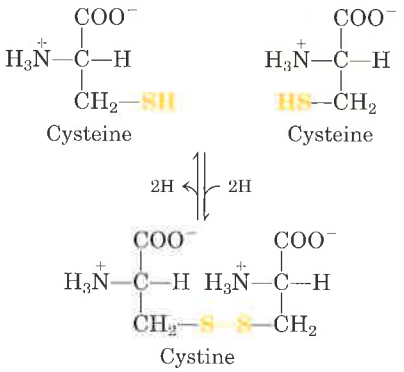

**Figure 5–7** Comparison of the light absorbance spectra of the aromatic amino acids at pH 6.0. The amino acids are present in equimolar amounts ($10^{-3}$ M) under identical conditions. The light absorbance of tryptophan is as much as fourfold higher than that of tyrosine. Phenylalanine absorbs less light than either tryptophan or tyrosine. Note that the absorbance maximum for tryptophan and tyrosine occurs near a wavelength of 280 nm.

***Aromatic R Groups*** **Phenylalanine, tyrosine,** and **tryptophan,** with their aromatic side chains (Fig. 5–6), are relatively nonpolar (hydrophobic). All can participate in hydrophobic interactions, which are particularly strong when the aromatic groups are stacked on one another. The hydroxyl group of tyrosine can form hydrogen bonds, and it acts as an important functional group in the activity of some enzymes. Tyrosine and tryptophan are significantly more polar than phenylalanine because of the tyrosine hydroxyl group and the nitrogen of the tryptophan indole ring.

Tryptophan and tyrosine, and to a lesser extent phenylalanine, absorb ultraviolet light (Fig. 5–7 and Box 5–1). This accounts for the characteristic strong absorbance of light by proteins at a wavelength of 280 nm, and is a property exploited by researchers in the characterization of proteins.

***Polar, Uncharged R Groups*** The R groups of these amino acids (Fig. 5–6) are more soluble in water, or hydrophilic, than those of the nonpolar amino acids, because they contain functional groups that form hydrogen bonds with water. This class of amino acids includes **serine, threonine, cysteine, methionine, asparagine,** and **glutamine.** The polarity of serine and threonine is contributed by their hydroxyl groups; that of cysteine and methionine by their sulfur atom; and that of asparagine and glutamine by their amide groups.

Asparagine and glutamine are the amides of two other amino acids also found in proteins, aspartate and glutamate, respectively, to which asparagine and glutamine are easily hydrolyzed by acid or base. Cysteine has an R group (a thiol group) that is approximately as acidic as the hydroxyl group of tyrosine. Cysteine requires special mention for another reason. It is readily oxidized to form a covalently linked dimeric amino acid called **cystine,** in which two cysteine molecules are joined by a disulfide bridge. Disulfide bridges of this kind occur in many proteins, stabilizing their structures.




***Negatively Charged (Acidic) R Groups*** The two amino acids having R groups with a net negative charge at pH 7.0 are **aspartate** and **glutamate,** each with a second carboxyl group (Fig. 5–6). These amino acids are the parent compounds of asparagine and glutamine, respectively.

***Positively Charged (Basic) R Groups*** The amino acids in which the R groups have a net positive charge at pH 7.0 are **lysine,** which has a second amino group at the $\epsilon$ position on its aliphatic chain; **arginine,** which has a positively charged guanidino group; and **histidine,** containing an imidazole group (Fig. 5–6). Histidine is the only standard amino acid having a side chain with a $pK_a$ near neutrality.

BOX 5–1

## Absorption of Light by Molecules: The Lambert–Beer Law

Measurement of light absorption is an important tool for analysis of many biological molecules. The fraction of the incident light absorbed by a solution at a given wavelength is related to the thickness of the absorbing layer (path length) and the concentration of the absorbing species. These two relationships are combined into the Lambert–Beer law, given in integrated form as

$$\log \frac{I_0}{I} = \epsilon c l$$

where $I_0$ is the intensity of the incident light, $I$ is the intensity of the transmitted light, $\epsilon$ is the molar absorption coefficient (in units of liters per mole-centimeter), $c$ the concentration of the absorbing species (in moles per liter), and $l$ the path length of the light-absorbing sample (in centimeters). The Lambert–Beer law assumes that the incident light is parallel and monochromatic and that the solvent and solute molecules are randomly oriented. The expression $\log (I_0/I)$ is called the absorbance, designated $A$.

It is important to note that each millimeter path length of absorbing solution in a 1.0 cm cell absorbs not a constant amount but a constant fraction of the incident light. However, with an absorbing layer of fixed path length, *the absorbance A is directly proportional to the concentration of the absorbing solute.*

The molar absorption coefficient varies with the nature of the absorbing compound, the solvent, the wavelength, and also with pH if the light-absorbing species is in equilibrium with another species having a different spectrum through gain or loss of protons.

In practice, absorbance measurements are usually made on a set of standard solutions of known concentration at a fixed wavelength. A sample of unknown concentration can then be compared with the resulting standard curve, as shown in Figure 1.

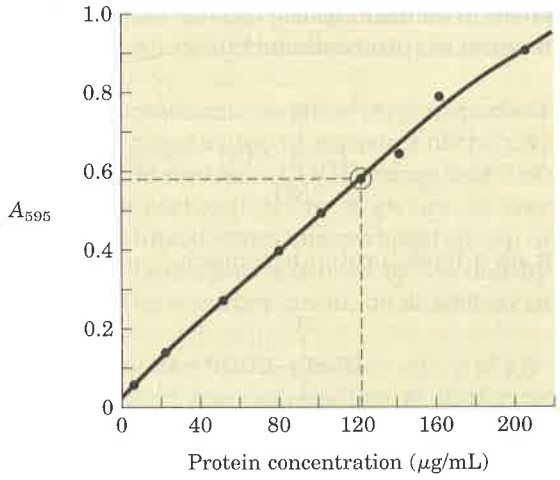
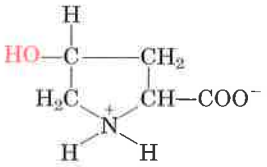


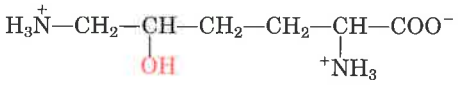

**Figure 1** Eight standard solutions containing known amounts of protein and one sample containing an unknown amount of protein were reacted with the Bradford reagent. This reagent contains a dye that shifts its absorption maximum to 595 nm when it binds amino acid residues. The $A_{595}$ (absorbance at 595 nm) of the standard samples was plotted against the protein concentration to create the standard curve, shown here. The $A_{595}$ of the unknown sample, 0.58, corresponds to a protein concentration of 122 μg/mL.

### Cells Also Contain Nonstandard Amino Acids

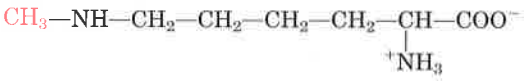
In addition to the 20 standard amino acids that are common in all proteins, other amino acids have been found as components of only certain types of proteins (Fig. 5–8a). Each of these is derived from one of the 20 standard amino acids, in a modification reaction that occurs after the standard amino acid has been inserted into a protein. Among the nonstandard amino acids are **4-hydroxyproline,** a derivative of proline, and **5-hydroxylysine;** the former is found in plant cell-wall proteins, and both are found in the fibrous protein collagen of connective tissues. ***N*-Methyllysine** is found in myosin, a contractile protein of muscle. Another important nonstandard amino acid is



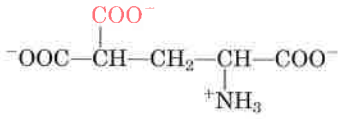
4-Hydroxyproline



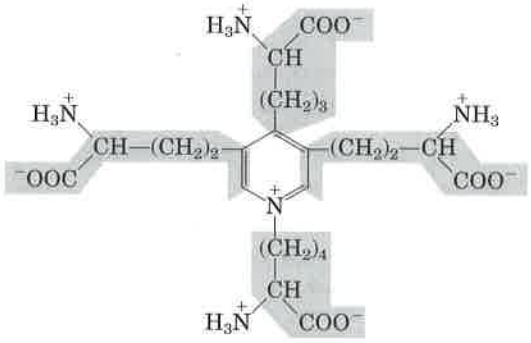
5-Hydroxylysine



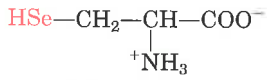
6-*N*-Methyllysine



γ-Carboxyglutamate



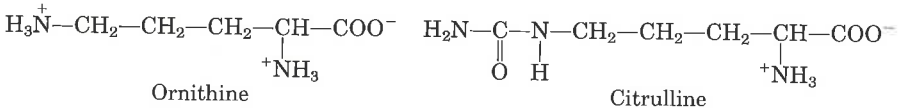
Desmosine



Selenocysteine

(a)

**Figure 5–8 (a)** Some nonstandard amino acids found in proteins; all are derived from standard amino acids. The extra functional groups are shown in red. Desmosine is formed from four residues of lysine, whose carbon backbones are shaded in gray. Selenocysteine is derived from serine. **(b)** Ornithine and citrulline are intermediates in the biosynthesis of arginine and in the urea cycle. Note that two systems are used to number carbons in the naming of these amino acids. The α, β, γ system used for γ-carboxyglutamate begins at the α carbon (see Fig. 5–2) and extends into the R group. The α-carboxyl group is not included. In contrast, the numbering system used to identify the modified carbon in 4-hydroxyproline, 5-hydroxylysine, and 6-*N*-methyllysine includes the α-carboxyl carbon, which is designated carbon 1 (or C-1).



Ornithine Citrulline

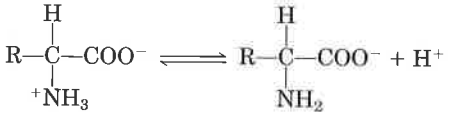
(b)

**γ-carboxyglutamate,** found in the blood-clotting protein prothrombin as well as in certain other proteins that bind $Ca^{2+}$ in their biological function. More complicated is the nonstandard amino acid **desmosine,** a derivative of four separate lysine residues, found in the fibrous protein elastin. **Selenocysteine** contains selenium rather than the oxygen of serine, and is found in glutathione peroxidase and a few other proteins.

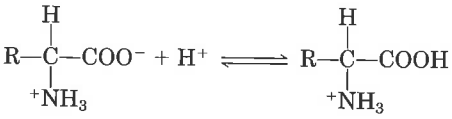
Some 300 additional amino acids have been found in cells and have a variety of functions but are not substituents of proteins. **Ornithine** and **citrulline** (Fig. 5–8b) deserve special note because they are key intermediates in the biosynthesis of arginine and in the urea cycle. These pathways are described in Chapters 21 and 17, respectively.

## Amino Acids Can Act as Acids and as Bases

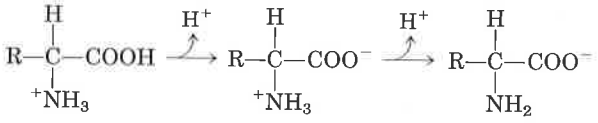
When a crystalline amino acid, such as alanine, is dissolved in water, it exists in solution as the dipolar ion, or zwitterion, which can act either as an acid (proton donor):



or as a base (proton acceptor):



Substances having this dual nature are **amphoteric** and are often called **ampholytes,** from "amphoteric electrolytes." A simple monoamino monocarboxylic α-amino acid, such as alanine, is actually a diprotic acid when it is fully protonated, that is, when both its carboxyl group and amino group have accepted protons. In this form it has two groups that can ionize to yield protons, as indicated in the following equation:



## Amino Acids Have Characteristic Titration Curves

Titration involves the gradual addition or removal of protons. Figure 5–9 shows the titration curve of the diprotic form of glycine. Each molecule of added base results in the net removal of one proton from

one molecule of amino acid. The plot has two distinct stages, each corresponding to the removal of one proton from glycine. Each of the two stages resembles in shape the titration curve of a monoprotic acid, such as acetic acid (see Fig. 4–10), and can be analyzed in the same way. At very low pH, the predominant ionic species of glycine is $^{+}H_3N{-}CH_2{-}COOH$, the fully protonated form. At the midpoint in the first stage of the titration, in which the —COOH group of glycine loses its proton, equimolar concentrations of proton-donor ($^{+}H_3N{-}CH_2{-}COOH$) and proton-acceptor ($^{+}H_3N{-}CH_2{-}COO^{-}$) species are present. At the midpoint of a titration (see Fig. 4–11), the pH is equal to the p$K_a$ of the protonated group being titrated. For glycine, the pH at the midpoint is 2.34, thus its —COOH group has a p$K_a$ of 2.34. [Recall that pH and p$K_a$ are simply convenient notations for proton concentration and the equilibrium constant for ionization, respectively (Chapter 4). The p$K_a$ is a measure of the tendency of a group to give up a proton, with that tendency decreasing tenfold as the p$K_a$ increases by one unit.] As the titration proceeds, another important point is reached at pH 5.97. Here there is a point of inflection, at which removal of the first proton is essentially complete, and removal of the second has just begun. At this pH the glycine is present largely as the dipolar ion $^{+}H_3N{-}CH_2{-}COO^{-}$. We shall return to the significance of this inflection point in the titration curve shortly.

The second stage of the titration corresponds to the removal of a proton from the $-NH_3^+$ group of glycine. The pH at the midpoint of this stage is 9.60, equal to the p$K_a$ for the $-NH_3^+$ group. The titration is complete at a pH of about 12, at which point the predominant form of glycine is $H_2N{-}CH_2{-}COO^{-}$.

From the titration curve of glycine we can derive several important pieces of information. First, it gives a quantitative measure of the p$K_a$ of each of the two ionizing groups, 2.34 for the —COOH group and 9.60 for the $-NH_3^+$ group. Note that the carboxyl group of glycine is over 100 times more acidic (more easily ionized) than the carboxyl group of acetic acid, which has a p$K_a$ of 4.76. This effect is caused by the nearby positively charged amino group on the $\alpha$-carbon atom, as described in Figure 5–10.

The second piece of information given by the titration curve of glycine (Fig. 5–9) is that this amino acid has *two* regions of buffering power (see Fig. 4–12). One of these is the relatively flat portion of the curve centered about the first p$K_a$ of 2.34, indicating that glycine is a good buffer near this pH. The other buffering zone extends for ~1.2 pH units centered around pH 9.60. Note also that glycine is not a good buffer at the pH of intracellular fluid or blood, about 7.4.

The Henderson–Hasselbalch equation (Chapter 4) can be used to calculate the proportions of proton-donor and proton-acceptor species of glycine required to make a buffer at a given pH within the buffering ranges of glycine; it also makes it possible to solve other kinds of buffer problems involving amino acids (see Box 4–2).

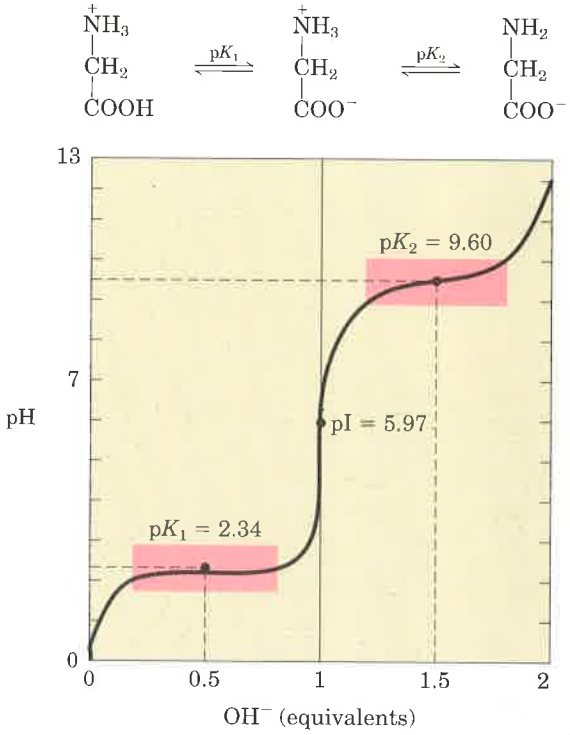
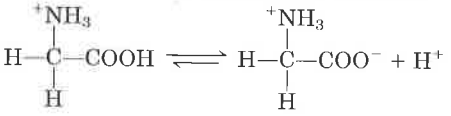



**Figure 5–9** The titration curve of 0.1 M glycine at 25 °C. The ionic species predominating at key points in the titration are shown above the graph. The shaded boxes, centered about p$K_1$ = 2.34 and p$K_2$ = 9.60, indicate the regions of greatest buffering power.

**Figure 5–10** **(a)** Interactions between the $\alpha$-amino and $\alpha$-carboxyl groups in an $\alpha$-amino acid. The nearby positive charge of the $-NH_3^+$ group makes ionization of the carboxyl group more likely (i.e., lowers the p$K_a$ for —COOH). This is due to a stabilizing interaction between opposite charges on the zwitterion and a repulsive interaction between the positive charges of the amino group and the departing proton. **(b)** The normal p$K_a$ for a carboxyl group is approximately 4.76, as for acetic acid.

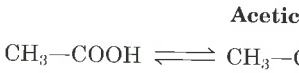
**$\alpha$-Amino acid (glycine)**




$pK_a = 2.34$

**(a)**

**Acetic acid**




$pK_a = 4.76$

**(b)**

### The Titration Curve Predicts the Electric Charge of Amino Acids

Another important piece of information derived from the titration curve of an amino acid is the relationship between its net electric charge and the pH of the solution. At pH 5.97, the point of inflection between the two stages in its titration curve, glycine is present as its dipolar form, fully ionized but with no *net* electric charge (Fig. 5–9). This characteristic pH is called the **isoelectric point** or **isoelectric pH,** designated **pI** or $pH_I$. For an amino acid such as glycine, which has no ionizable group in the side chain, the isoelectric point is the arithmetic mean of the two $pK_a$ values:

$$pI = \frac{1}{2}(pK_1 + pK_2)$$

which in the case of glycine is

$$pI = \frac{1}{2}(2.34 + 9.60) = 5.97$$

As is evident in Figure 5–9, glycine has a net negative charge at any pH above its pI and will thus move toward the positive electrode (the anode) when placed in an electric field. At any pH below its pI, glycine has a net positive charge and will move toward the negative electrode, the cathode. The farther the pH of a glycine solution is from its isoelectric point, the greater the net electric charge of the population of glycine molecules. At pH 1.0, for example, glycine exists entirely as the form $^+H_3N—CH_2—COOH$, with a net positive charge of 1.0. At pH 2.34, where there is an equal mixture of $^+H_3N—CH_2—COOH$ and $^+H_3N—CH_2—COO^-$, the average or net positive charge is 0.5. The sign and the magnitude of the net charge of any amino acid at any pH can be predicted in the same way.

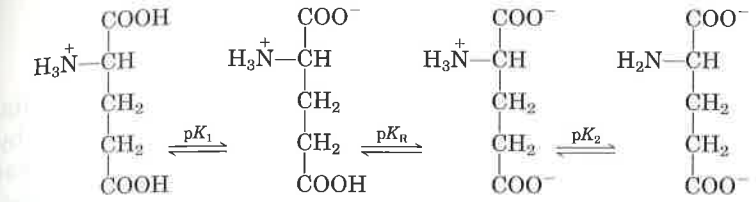
This information has practical importance. For a solution containing a mixture of amino acids, the different amino acids can be separated on the basis of the direction and relative rate of their migration when placed in an electric field at a known pH.

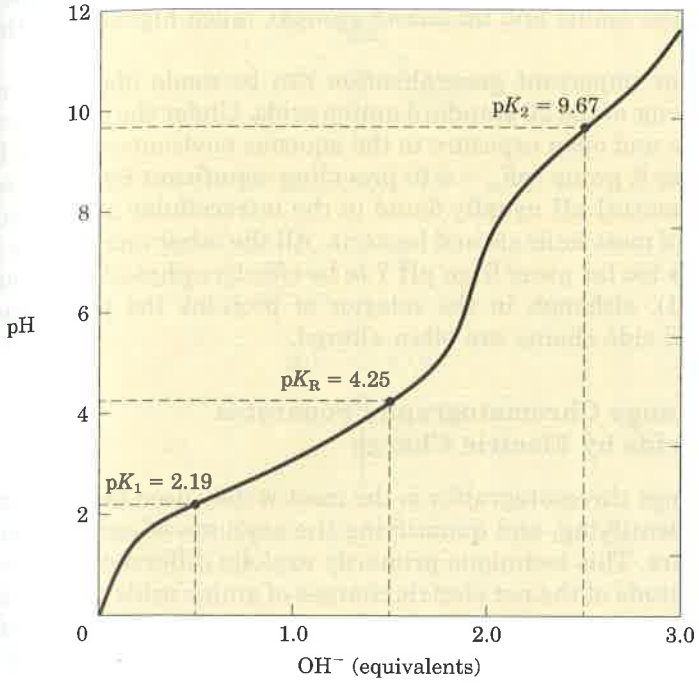
### Amino Acids Differ in Their Acid–Base Properties

The shared properties of many amino acids permit some simplifying generalizations about the acid–base behavior of different classes of amino acids.

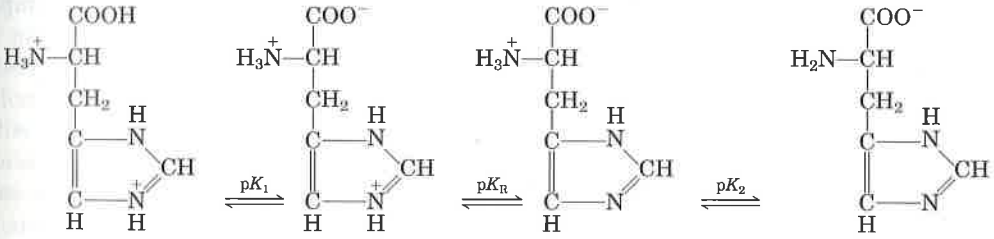
All amino acids with a single $\alpha$-amino group, a single $\alpha$-carboxyl group, and an R group that does not ionize have titration curves resembling that of glycine (Fig. 5–9). This group of amino acids is characterized by having very similar, although not identical, values for $pK_1$ (the $pK$ of the —COOH group) in the range of 1.8 to 2.4 and for $pK_2$ (of the $—NH_3^+$ group) in the range of 8.8 to 11.0 (Table 5–1).

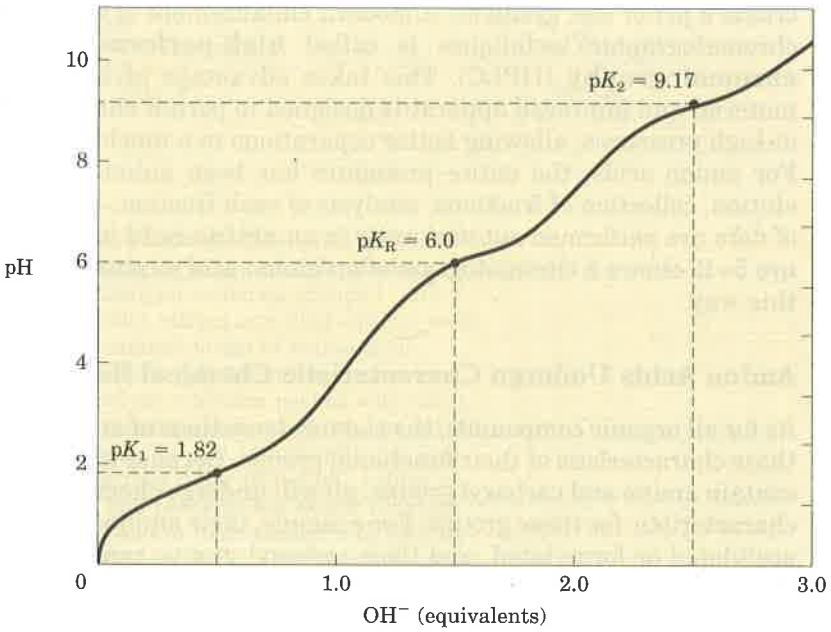
Amino acids with an ionizable R group (Table 5–1) have more complex titration curves with *three* stages corresponding to the three possible ionization steps; thus they have three $pK_a$ values. The third stage for the titration of the ionizable R group merges to some extent with the others. The titration curves of two representatives of this group, glutamate and histidine, are shown in Figure 5–11. The isoelectric points of amino acids in this class reflect the type of ionizing R groups present. For example, glutamate has a pI of 3.22, considerably lower than that of glycine. This is a result of the presence of two carboxyl




**(a)**




**(b)**

**Figure 5–11** The titration curves of **(a)** glutamate and **(b)** histidine. The $pK_a$ of the R group is designated $pK_R$.

groups which, at the average of their $pK_a$ values (3.22), contribute a net negative charge of −1 that balances the +1 contributed by the amino group. Similarly, the pI of histidine, with two groups that are positively charged when protonated, is 7.59 (the average of the $pK_a$ values of the amino and imidazole groups), much higher than that of glycine.

Another important generalization can be made about the acid–base behavior of the 20 standard amino acids. Under the general condition of free and open exposure to the aqueous environment, only histidine has an R group ($pK_a$ = 6.0) providing significant buffering power near the neutral pH usually found in the intracellular and intercellular fluids of most animals and bacteria. All the other amino acids have $pK_a$ values too far away from pH 7 to be effective physiological buffers (Table 5–1), although in the interior of proteins the $pK_a$ values of amino acid side chains are often altered.

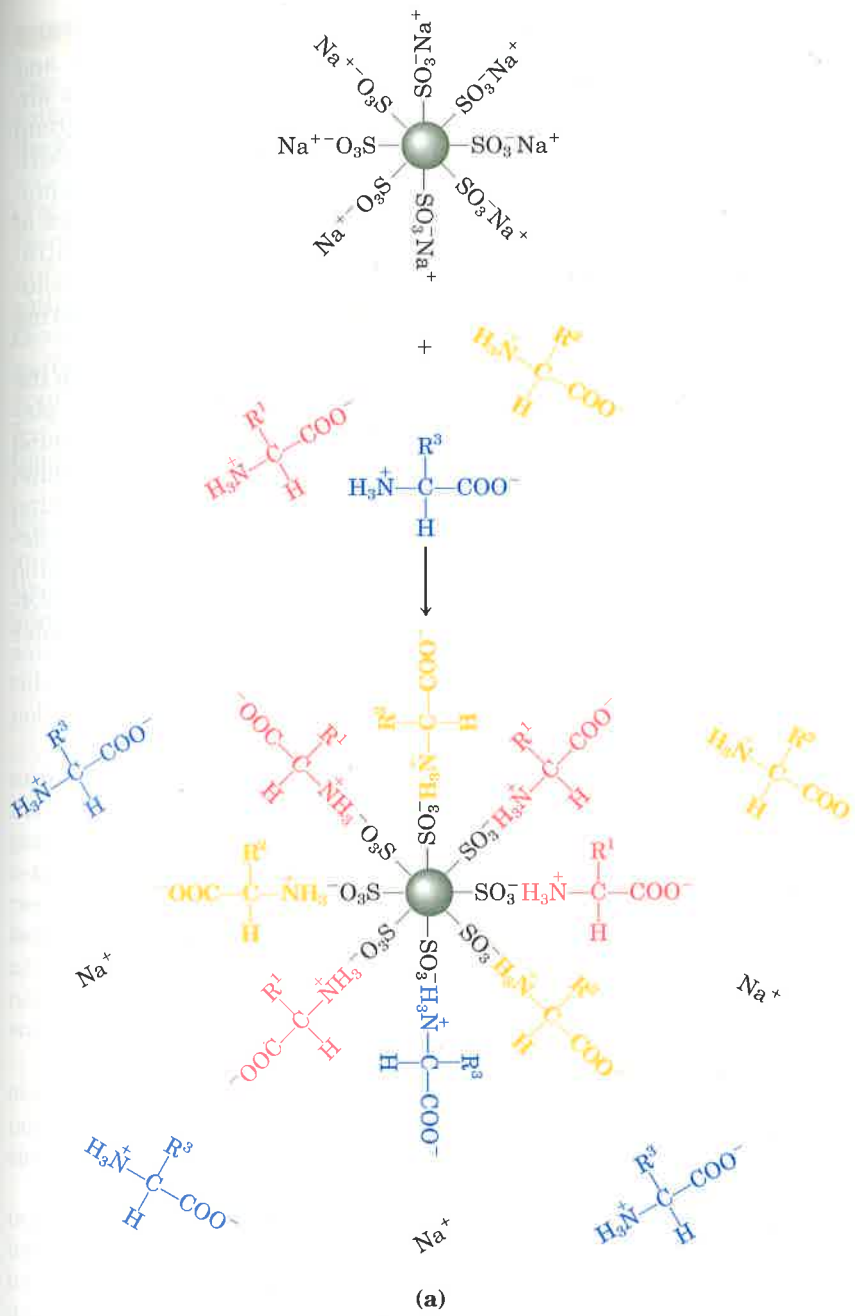
### Ion-Exchange Chromatography Separates Amino Acids by Electric Charge

Ion-exchange chromatography is the most widely used method for separating, identifying, and quantifying the amounts of each amino acid in a mixture. This technique primarily exploits differences in the sign and magnitude of the net electric charges of amino acids at a given pH, which are predictable from their $pK_a$ values or their titration curves.

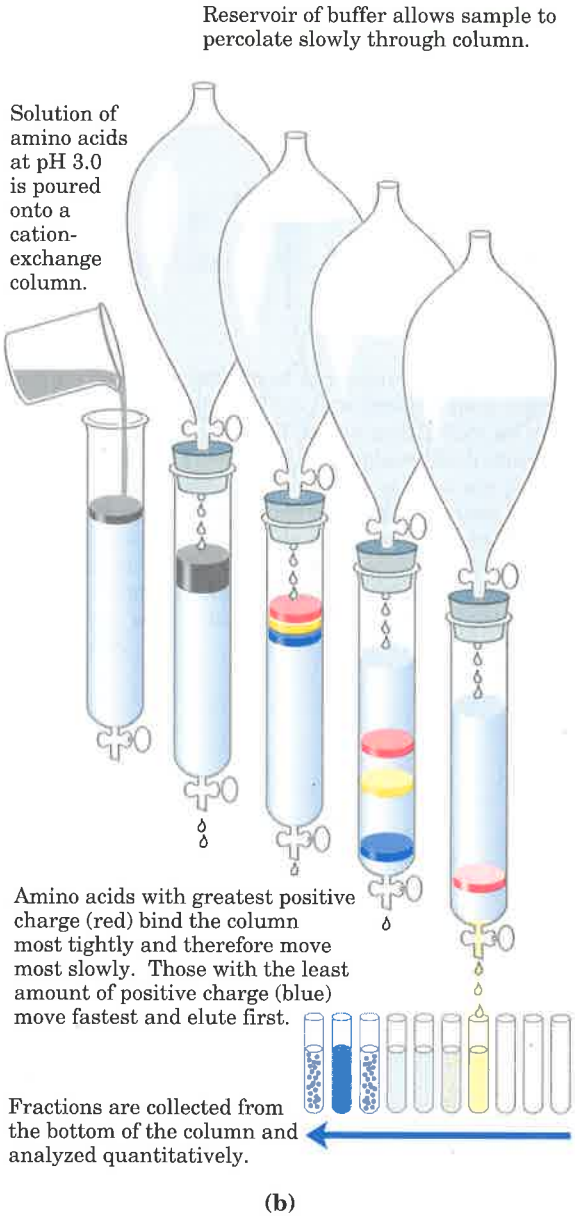
The chromatographic column consists of a long tube filled with particles of a synthetic resin containing fixed charged groups; those with fixed anionic groups are called **cation-exchange** resins and those with fixed cationic groups, **anion-exchange** resins. A simple form of ion-exchange chromatography on a cation-exchange resin is described in Figure 5–12. The affinity of each amino acid for the resin is affected by pH (which determines the ionization state of the molecule) and the concentration of other salt ions that may compete with the resin by associating with the amino acid. Separation of amino acids can therefore be optimized by gradually changing the pH and/or salt concentration of the solution being passed through the column so as to create a pH or salt gradient. A modern enhancement of this and other chromatographic techniques is called **high-performance liquid chromatography** (HPLC). This takes advantage of stronger resin material and improved apparatus designed to permit chromatography at high pressures, allowing better separations in a much shorter time. For amino acids, the entire procedure has been automated, so that elution, collection of fractions, analysis of each fraction, and recording of data are performed automatically in an **amino-acid analyzer.** Figure 5–13 shows a chromatogram of an amino acid mixture analyzed in this way.

### Amino Acids Undergo Characteristic Chemical Reactions

As for all organic compounds, the chemical reactions of amino acids are those characteristic of their functional groups. Because all amino acids contain amino and carboxyl groups, all will undergo chemical reactions characteristic for these groups. For example, their amino groups can be acetylated or formylated, and their carboxyl groups can be esterified. We will not examine all such organic reactions of amino acids, but several widely used reactions are noteworthy because they greatly simplify the detection, measurement, and identification of amino acids.



(a)




(b)

**Figure 5–12** Ion-exchange chromatography. An example of a cation-exchange resin is presented. **(a)** Negatively charged sulfonate groups ($-SO_3^-$) on the resin surface attract and bind cations, such as $H^+$, $Na^+$, or cationic forms of amino acids. **(b)** An acidic solution (pH 3.0) of the amino acid mixture is poured on a column packed with resin and allowed to percolate through slowly. At pH 3.0 the amino acids are largely cations with net positive charges, but they differ in the $pK_a$ values of their R groups, and hence in the extent to which they are ionized and in their tendency to bind to the anionic resin. As a result, they move through the column at different rates.

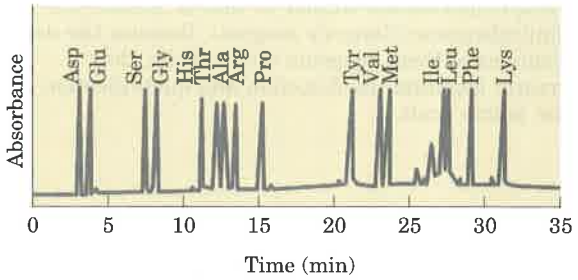



**Figure 5–13** Automatically recorded high-performance liquid chromatographic analysis of amino acids on a cation-exchange resin. The area under each peak on the chromatogram is proportional to the amount of each amino acid in the mixture.

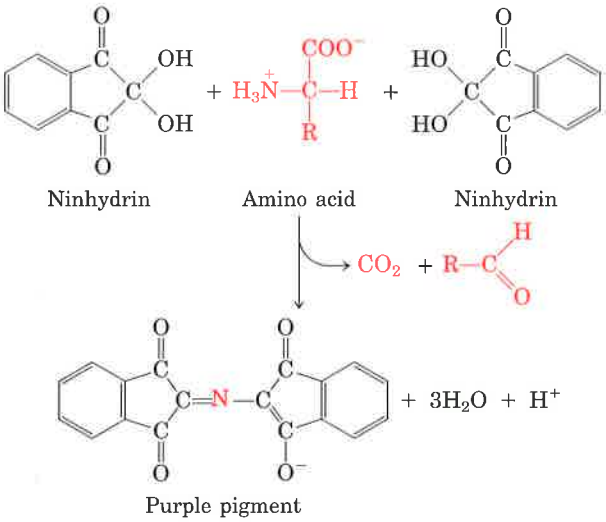
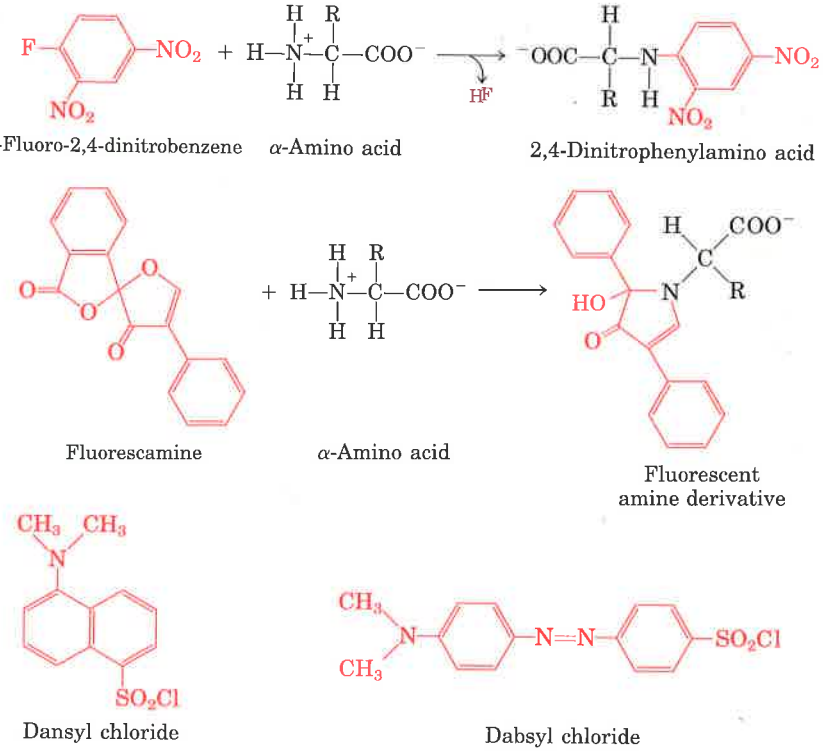



**Figure 5–14** Reagents that react with the α-amino group of amino acids. The reactions producing 2,4-dinitrophenyl and fluorescamine derivatives are illustrated. The reactions of dansyl chloride and dabsyl chloride are similar to that of 1-fluoro-2,4-dinitrobenzene (Sanger's reagent). Because the derivatives of these reagents absorb light, they greatly facilitate the detection and quantification of the amino acids.

One of the most important, technically and historically, is the ninhydrin reaction, which has been used for many years to detect and quantify microgram amounts of amino acids. When amino acids are heated with excess ninhydrin, all those having a free α-amino group yield a purple product. Proline, in which the α-amino group is substituted (forming an imino group), yields a yellow product. Under appropriate conditions the intensity of color produced (optical absorbance of the solution; see Box 5–1) is proportional to the amino acid concentration. Comparing the absorbance to that of appropriate standard solutions is an accurate and technically simple method for measuring amino acid concentration.

Several other convenient reagents are available that react with the α-amino group to form colored or fluorescent derivatives. Unlike ninhydrin, these have the advantage that the intact R group of the amino acid remains part of the product, so that derivatives of different amino acids can be distinguished. Fluorescamine reacts rapidly with amino acids and provides great sensitivity, yielding a highly fluorescent derivative that permits the detection of nanogram quantities of amino acids (Fig. 5–14). Dabsyl chloride, dansyl chloride, and 1-fluoro-2,4-dinitrobenzene (Fig. 5–14) yield derivatives that are stable under harsh conditions such as those used in the hydrolysis of proteins.




## Peptides

We now turn to polymers of amino acids, the **peptides.** Biologically occurring peptides range in size from small molecules containing only two or three amino acids to macromolecules containing thousands of amino acids. The focus here is on the structure and chemical properties of the smaller peptides, providing a prelude to the discussion of the large peptides called proteins in the next two chapters.

## Peptides Are Chains of Amino Acids

Two amino acid molecules can be covalently joined through a substituted amide linkage, termed a **peptide bond,** to yield a dipeptide. Such a linkage is formed by removal of the elements of water from the $\alpha$-carboxyl group of one amino acid and the $\alpha$-amino group of another (Fig. 5–15). Peptide-bond formation is an example of a condensation reaction, a common class of reaction in living cells. Note that as shown in Figure 5–15, this reaction has an equilibrium that favors reactants rather than products. To make the reaction thermodynamically more favorable, the carboxyl group must be chemically modified or activated so that the hydroxyl group can be more readily eliminated. A chemical approach to this problem is outlined at the end of this chapter (see Box 5–2). The biological approach to peptide bond formation is a major topic of Chapter 26.

Three amino acids can be joined by two peptide bonds to form a tripeptide; similarly, amino acids can be linked to form tetrapeptides and pentapeptides. When a few amino acids are joined in this fashion, the structure is called an **oligopeptide.** When many amino acids are joined, the product is called a **polypeptide.** Proteins may have thousands of amino acid units. Although the terms "protein" and "polypeptide" are sometimes used interchangeably, molecules referred to as polypeptides generally have molecular weights below 10,000.

Figure 5–16 shows the structure of a pentapeptide. The amino acid units in a peptide are often called **residues** (each has lost a hydrogen atom from its amino group and a hydroxyl moiety from its carboxyl group). The amino acid residue at that end of a peptide having a free $\alpha$-amino group is the **amino-terminal** (or N-terminal) residue; the residue at the other end, which has a free carboxyl group, is the **carboxyl-terminal** (C-terminal) residue. By convention, short peptides are named from the sequence of their constituent amino acids, beginning at the left with the amino-terminal residue and proceeding toward the carboxyl terminus at the right (Fig. 5–16).

Although hydrolysis of peptide bonds is an exergonic reaction, it occurs slowly because of its high activation energy. As a result, the peptide bonds in proteins are quite stable under most intracellular conditions.

The peptide bond is the single most important covalent bond linking amino acids in peptides and proteins. The only other type of covalent bond that occurs frequently enough to deserve special mention here is the disulfide bond sometimes formed between two cysteine residues. Disulfide bonds play a special role in the structure of many proteins, particularly those that function extracellularly, such as the hormone insulin and the immunoglobulins or antibodies.

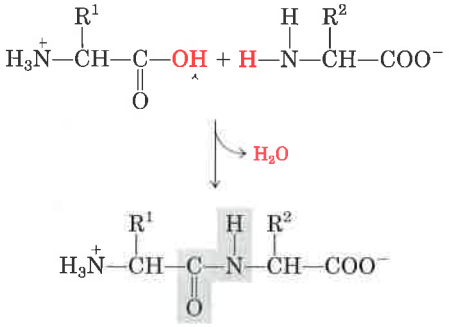



**Figure 5–15** Formation of a peptide bond (shaded in gray) in a dipeptide. This is a condensation reaction. The $\alpha$-amino group of amino acid 2 acts as a nucleophile (see Table 3–6) to displace the hydroxyl group of amino acid 1 (red). Amino groups are good nucleophiles, but the hydroxyl group is a poor leaving group and is not readily displaced. At physiological pH the reaction as shown does not occur to any appreciable extent. Peptide bond formation is endergonic, with a free-energy change of about +21 kJ/mol.

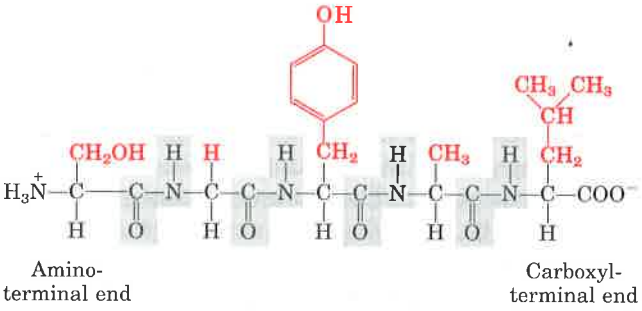



**Figure 5–16** Structure of the pentapeptide seryl-glycyltyrosinylalanylleucine, or Ser–Gly–Tyr–Ala–Leu. Peptides are named beginning with the amino-terminal residue, which by convention is placed at the left. The peptide bonds are shown shaded in gray, the R groups in red.

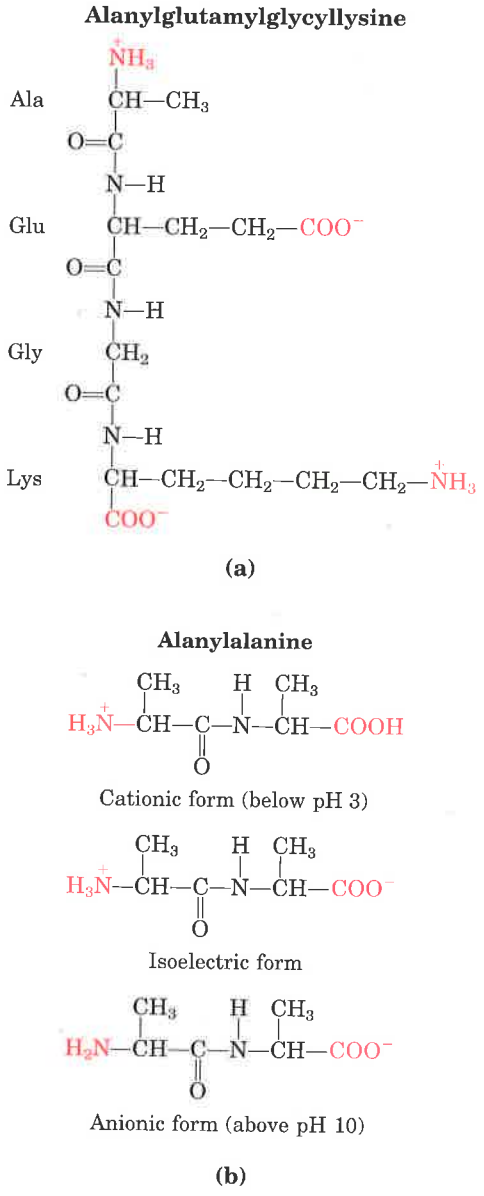



**Figure 5–17** Ionization and electric charge of peptides. The groups ionized at pH 7.0 are in red. **(a)** A tetrapeptide with two ionizable R groups. **(b)** The cationic, isoelectric, and anionic forms of a dipeptide lacking ionizable R groups.

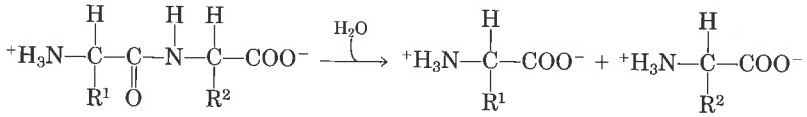
## Peptides Can Be Distinguished by Their Ionization Behavior

Peptides contain only one free $\alpha$-amino group and one free $\alpha$-carboxyl group (Fig. 5–17). These groups ionize as they do in simple amino acids, although the ionization constants are different because the oppositely charged group is absent from the $\alpha$ carbon. The $\alpha$-amino and $\alpha$-carboxyl groups of all other constituent amino acids are covalently joined in the form of peptide bonds, which do not ionize and thus do not contribute to the total acid–base behavior of peptides. However, the R groups of some amino acids can ionize (Table 5–1), and in a peptide these contribute to the overall acid–base properties (Fig. 5–17). Thus the acid–base behavior of a peptide can be predicted from its single free $\alpha$-amino and $\alpha$-carboxyl groups and the nature and number of its ionizable R groups. Like free amino acids, peptides have characteristic titration curves and a characteristic isoelectric pH at which they do not move in an electric field. These properties are exploited in some of the techniques used to separate peptides and proteins (Chapter 6).

## Peptides Undergo Characteristic Chemical Reactions

Like other organic molecules, peptides undergo chemical reactions that are characteristic of their functional groups: the free amino and carboxyl groups and the R groups.

Peptide bonds can be hydrolyzed by boiling with either strong acid (typically 6 M HCl) or base to yield the constituent amino acids.



Hydrolysis of peptide bonds in this manner is a necessary step in determining the amino acid composition of proteins. The reagents shown in Figure 5–14 label only free amino groups: those of the amino-terminal residue and the R groups of any lysines present. If dabsyl chloride, dansyl chloride, or 1-fluoro-2,4-dinitrobenzene is used before acid hydrolysis of the peptide, the amino-terminal residue can be separated and identified (Fig. 5–18).

Peptide bonds can also be hydrolyzed by certain enzymes called **proteases.** Proteolytic (protein-cleaving) enzymes are found in all cells and tissues, where they degrade unneeded or damaged proteins or aid in the digestion of food.

## Some Small Polypeptides Have Biological Activity

Much of the material in the chapters to follow will revolve around the activities of proteins with molecular weights measured in the tens and even hundreds of thousands. Not all polypeptides are so large, however. There are many naturally occurring small polypeptides and oligopeptides, some of which have important biological activities and exert their effects at very low concentrations. For example, a number of vertebrate hormones (intercellular chemical messengers) (Chapter 22) are small polypeptides. The hormone insulin contains two polypeptide chains, one having 30 amino acid residues and the other 21. Other polypeptide hormones include glucagon, a pancreatic hormone of 29 residues that opposes the action of insulin, and corticotropin, a 39-

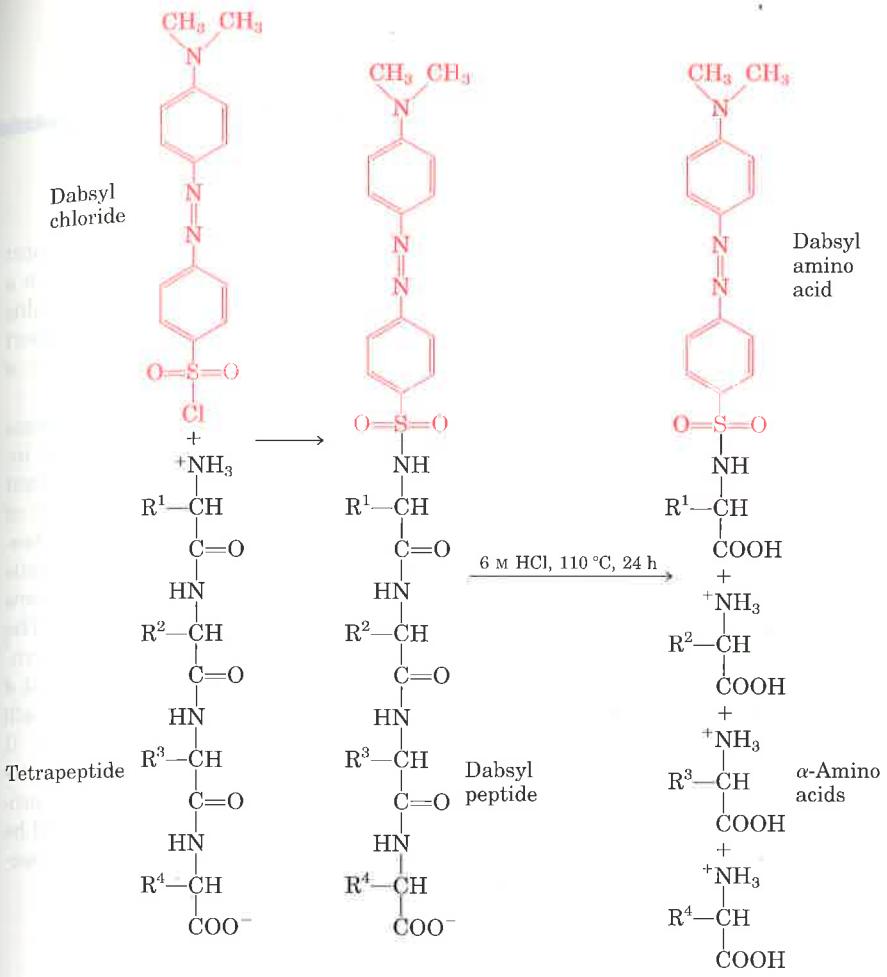
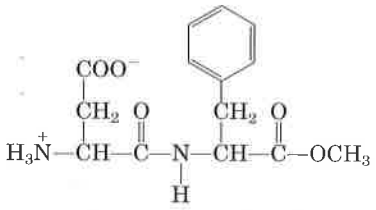



**Figure 5–18** The amino-terminal residue of a tetrapeptide can be identified by labeling it with dabsyl chloride, then hydrolyzing the peptide bonds in strong acid. The result is a mixture of amino acids of which only the amino-terminal amino acid (and lysine) is labeled.

residue hormone of the anterior pituitary gland that stimulates the adrenal cortex.

Some biologically important peptides have only a few amino acid residues. That small peptides can have large biological effects is readily illustrated by the activity of the commercially synthesized dipeptide, L-aspartylphenylalanyl methyl ester. This compound is an artificial sweetener better known as aspartame or NutraSweet®:



L-Aspartyl-L-phenylalanyl methyl ester (aspartame)

Among naturally occurring small peptides are hormones such as oxytocin (nine amino acid residues), which is secreted by the posterior pituitary and stimulates uterine contractions; bradykinin (nine residues), which inhibits inflammation of tissues; and thyrotropin-releasing factor (three residues), which is formed in the hypothalamus and stimulates the release of another hormone, thyrotropin, from the anterior pituitary gland (Fig. 5–19). Also noteworthy among short peptides are the enkephalins, compounds formed in the central nervous system

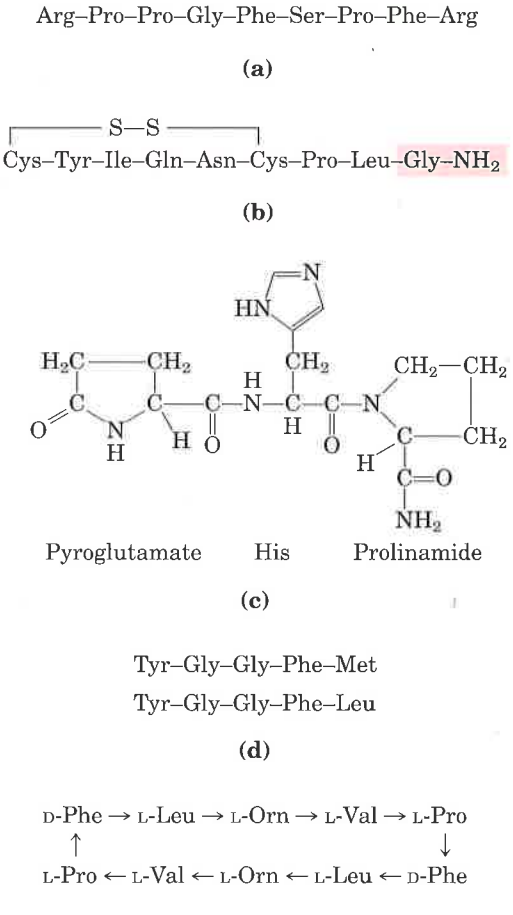



**Figure 5–19** Some naturally occurring peptides with intense biological activity. The amino-terminal residues are at the left end. **(a)** Bradykinin, a hormonelike peptide that inhibits inflammatory reactions. **(b)** Oxytocin, formed by the posterior pituitary gland. The shaded portion is a residue of glycinamide ($H_2N—CH_2—CONH_2$). **(c)** Thyrotropin-releasing factor, formed by the hypothalamus. **(d)** Two enkephalins, brain peptides that affect the perception of pain. **(e)** Gramicidin S, an antibiotic produced by the bacterium *Bacillus brevis*. The arrows indicate the direction from the amino toward the carboxyl end of each residue. The peptide has no termini because it is circular. Orn is the symbol for ornithine, an amino acid that generally does not occur in proteins. Note that gramicidin S contains two residues of a D-amino acid (D-phenylalanine).

## BOX 5–2 Chemical Synthesis of Peptides and Small Proteins

Many peptides are potentially useful as pharmacological reagents, and their synthesis is of considerable commercial importance. There are three ways to obtain a peptide: (1) purification from tissue, a task often made difficult by the vanishingly low concentrations of some peptides; (2) genetic engineering; or (3) direct chemical synthesis. Powerful techniques now make direct chemical synthesis an attractive option in many cases. In addition to commercial applications, the synthesis of specific peptide portions of larger proteins is an increasingly important tool for the study of protein structure and function.

The complexity of proteins makes the traditional synthetic approaches of organic chemistry impractical for peptides with more than four or five amino acids. One problem is the difficulty of purifying the product after each step, because the chemical properties of the peptide change each time a new amino acid is added.

The major breakthrough in this technology was provided by R. Bruce Merrifield. His innovation involved synthesizing a peptide while keeping it attached at one end to a solid support. The support is an insoluble polymer (resin) contained within a column, similar to that used for chromatographic procedures. The peptide is built up on this support one amino acid at a time using a standard set of reactions in a repeating cycle (Fig. 1).

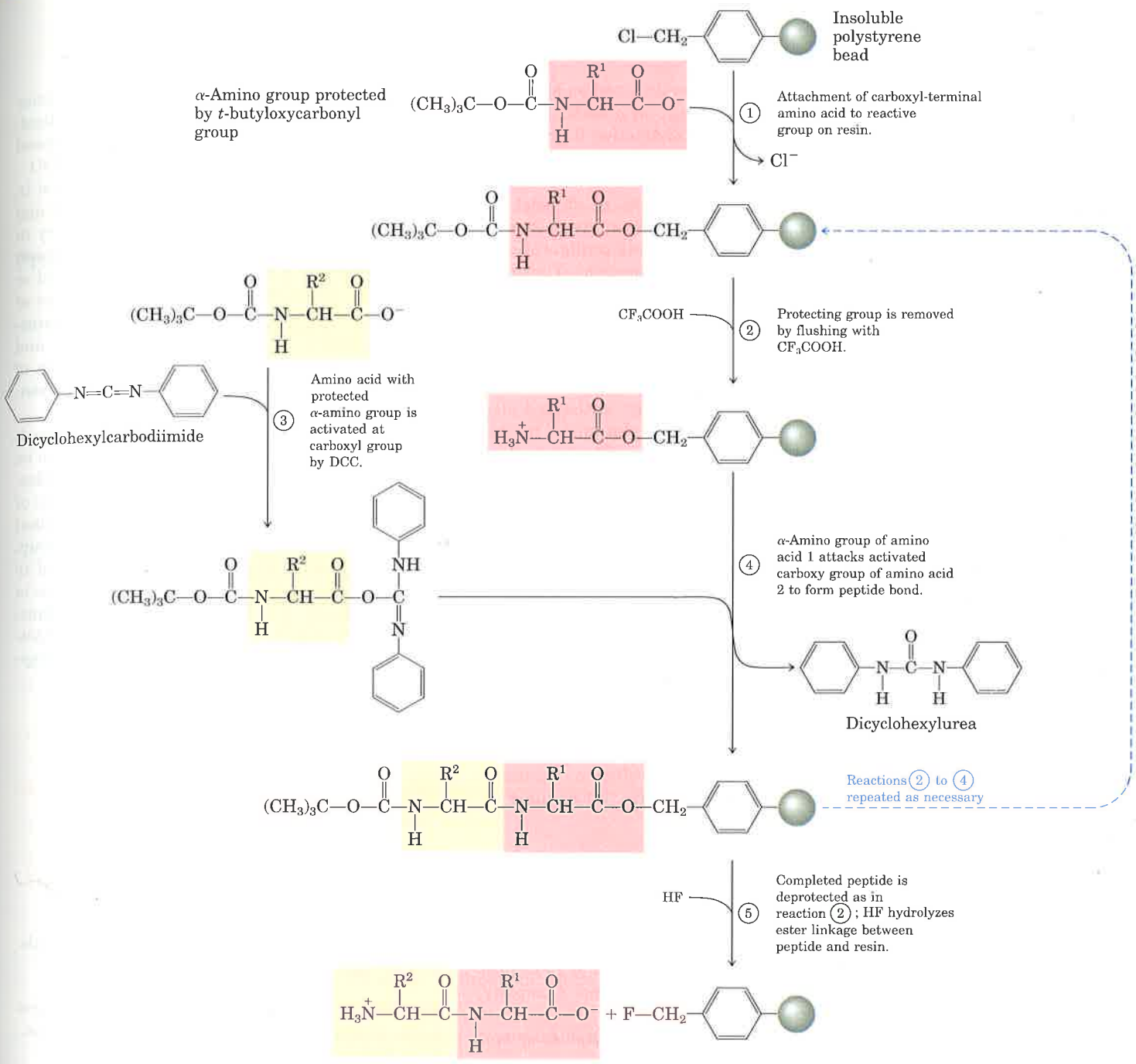
The technology for chemical peptide synthesis has been automated, and several commercial instruments are now available. The most important limitation of the process involves the efficiency of each amino acid addition, as can be seen by calculating the overall yields of peptides of various lengths when the yield for addition of each new amino acid is 96.0 versus 99.8% (Table 1). The chemistry has been optimized to permit the synthesis of proteins 100 amino acids long in about 4 days in reasonable yield. A very similar approach is used to synthesize nucleic acids (Fig. 12–38). It is worth noting that this technology, impressive as it is, still pales when compared with biological processes. The same 100 amino-acid protein would be synthesized with exquisite fidelity in about 5 seconds in a bacterial cell.

**Table 1 Effect of stepwise yield on overall yields in peptide synthesis**

| Number of residues in the final polypeptide | Overall yields of final peptide (%) when the yield of each step is: 96.0% | 99.8% |
|---|---|---|
| 11 | 66 | 98 |
| 21 | 44 | 96 |
| 31 | 29 | 94 |
| 51 | 13 | 90 |
| 100 | 1.7 | 82 |

**Figure 1** Chemical synthesis of a peptide on a solid support. Reactions ② through ④ are necessary for the formation of each peptide bond.

that bind to receptors in certain cells of the brain and induce analgesia (deadening of pain sensations). Enkephalins represent one of the body's own mechanisms for control of pain. The enkephalin receptors also bind morphine, heroin, and other addicting opiate drugs (although these are not peptides). Some extremely toxic mushroom poisons, such as amanitin, are also peptides, as are many antibiotics.




A growing number of small peptides are proving to be important commercially as pharmaceutical reagents. Unfortunately, they are often present in exceedingly small amounts and hence are hard to purify. For these and other reasons, the chemical synthesis of peptides has become one of the major technologies associated with biochemistry (Box 5–2).

## Summary

The 20 amino acids commonly found as hydrolysis products of proteins contain an $\alpha$-carboxyl group, an $\alpha$-amino group, and a distinctive R group substituted on the $\alpha$-carbon atom. The $\alpha$-carbon atom of the amino acids (except glycine) is asymmetric, and thus amino acids can exist in at least two stereoisomeric forms. Only the L stereoisomers, which are related to the absolute configuration of L-glyceraldehyde, are found in proteins. The amino acids are classified on the basis of the polarity of their R groups. The nonpolar, aliphatic class includes alanine, glycine, isoleucine, leucine, proline, and valine. Phenylalanine, tryptophan, and tyrosine have aromatic side chains and are also relatively hydrophobic. The polar, uncharged class includes asparagine, cysteine, glutamine, methionine, serine, and threonine. The negatively charged (acidic) amino acids are aspartate and glutamate; the positively charge (basic) ones are arginine, histidine, and lysine. There are also a large number of nonstandard amino acids that occur in some proteins (as a result of the modification of standard amino acids) or as free metabolites in cells.

Monoamino monocarboxylic amino acids are diprotic acids ($^+H_3NCH(R)COOH$) at low pH. As the pH is raised to about 6, near the isoelectric point, the proton is lost from the carboxyl group to form the dipolar or zwitterionic species $^+H_3NCH(R)COO^-$, which is electrically neutral. Further increase in pH causes loss of the second proton, to yield the ionic species $H_2NCH(R)COO^-$. Amino acids with ionizable R groups may exist in additional ionic species, depending on the pH and the $pK_a$ of the R group. Thus amino acids vary in their acid–base properties. Amino acids form colored derivatives with ninhydrin. Other colored or fluorescent derivatives are formed in reactions of the $\alpha$-amino group of amino acids with fluorescamine, dansyl chloride, dabsyl chloride, and 1-fluoro-2,4-dinitrobenzene. Complex mixtures of amino acids can be separated and identified by ion-exchange chromatography or HPLC.

Amino acids can be joined covalently through peptide bonds to form peptides, which can also be formed by incomplete hydrolysis of polypeptides. The acid–base behavior and chemical reactions of a peptide are functions of its amino-terminal amino group, its carboxyl-terminal carboxyl group, and its R groups. Peptides can be hydrolyzed to yield free amino acids. Some peptides occur free in cells and tissues and have specific biological functions. These include some hormones and antibiotics, as well as other peptides with powerful biological activity.

## Further Reading

### *General*

**Cantor, C.R. & Schimmel, P.R.** (1980) *Biophysical Chemistry,* Part I: *The Conformation of Biological Macromolecules,* W.H. Freeman and Company, San Francisco.

*Excellent textbook outlining the properties of biological macromolecules and their monomeric subunits.*

**Creighton, T.E.** (1984) *Proteins: Structures and Molecular Properties,* W.H. Freeman and Company, New York.

*Very useful general source.*

**Dickerson, R.E. & Geis, I.** (1983) *Proteins: Structure, Function, and Evolution,* 2nd edn, The Benjamin/Cummings Publishing Company, Menlo Park, CA.

*Beautifully illustrated and interesting account.*

### *Amino Acids*

**Corrigan, J.J.** (1969) D-Amino acids in animals. *Science* **169,** 142–148.

**Meister, A.** (1965) *Biochemistry of the Amino Acids,* 2nd edn, Vols. 1 and 2, Academic Press, Inc., New York.

*Encyclopedic treatment of the properties, occurrence, and metabolism of amino acids.*

**Montgomery, R. & Swenson, C.A.** (1976) *Quantitative Problems in the Biochemical Sciences,* 2nd edn, W.H. Freeman and Company, New York.

**Segel, I.H.** (1976) *Biochemical Calculations,* 2nd edn, John Wiley & Sons, New York.

***Peptides***

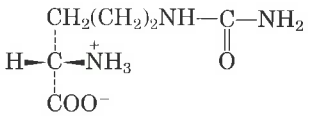
**Haschemeyer, R.H. & Haschemeyer, A.E.V.** (1973) *Proteins: A Guide to Study by Physical and Chemical Methods,* John Wiley & Sons, New York.

**Merrifield, B.** (1986) Solid phase synthesis. *Science* **232,** 341–347.

**Smith, L.M.** (1988) Automated synthesis and sequence analysis of biological macromolecules. *Analyt. Chem.* **60,** 381A–390A.

## Problems

**1.** *Absolute Configuration of Citrulline* Is citrulline isolated from watermelons (shown below) a D- or L-amino acid? Explain.



**2.** *Relation between the Structures and Chemical Properties of the Amino Acids* The structures and chemical properties of the amino acids are crucial to understanding how proteins carry out their biological functions. The structures of the side chains of 16 amino acids are given below. Name the amino acid that contains each structure and match the R group with the most appropriate description of its properties, (a) to (m). Some of the descriptions may be used more than once.

(a) Small polar R group containing a hydroxyl group; this amino acid is important in the active site of some enzymes.

(b) Provides the least amount of steric hindrance.

(c) R group has $pK_a \approx 10.5$, making it positively charged at physiological pH.

(d) Sulfur-containing R group; neutral at any pH.

(e) Aromatic R group, hydrophobic in nature and neutral at any pH.

(f) Saturated hydrocarbon, important in hydrophobic interactions.

(g) The only amino acid having an ionizing R group with a $pK_a$ near 7; it is an important group in the active site of some enzymes.

(h) The only amino acid having a substituted $\alpha$-amino group; it influences protein folding by forcing a bend in the chain.

(i) R group has a $pK_a$ near 4 and thus is negatively charged at pH 7.

(j) An aromatic R group capable of forming hydrogen bonds; it has a $pK_a$ near 10.

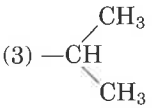
(k) Forms disulfide cross-links between polypeptide chains; the $pK_a$ of its functional group is about 10.

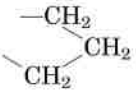
(l) R group with $pK_a \approx 12$, making it positively charged at physiological pH.

(m) When this polar but uncharged R group is hydrolyzed, the amino acid is converted into another amino acid having a negatively charged R group at pH near 7.

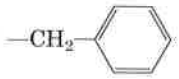
(1) —H

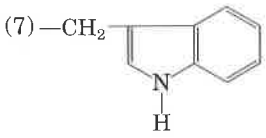
(2) —$CH_3$

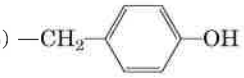
(3) 

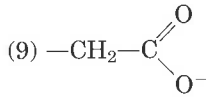
(4) 

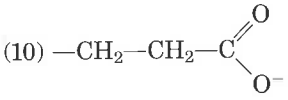
(5) —$CH_2OH$

(6) 

(7) 

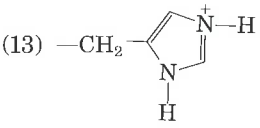
(8) 

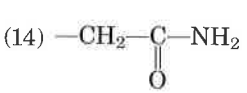
(9) 

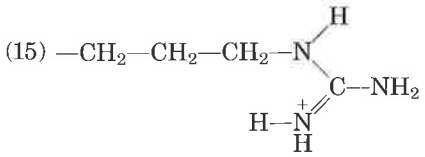
(10) 

(11) —$CH_2$—$CH_2$—S—$CH_3$

(12) —$CH_2$—SH

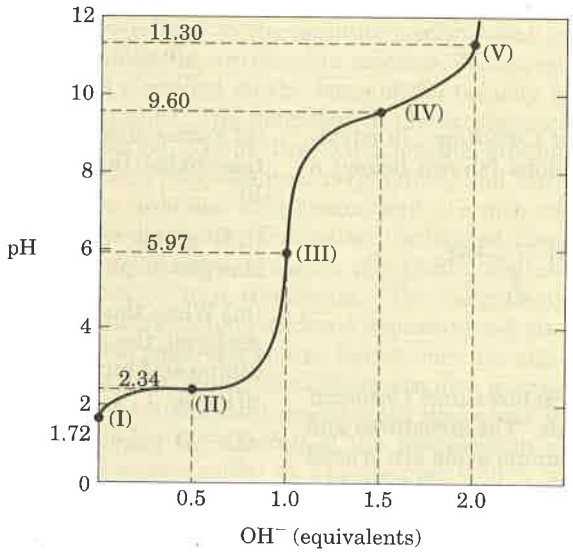
(13) 

(14) 

(15) 

(16) —$CH_2$—$CH_2$—$CH_2$—$CH_2$—$\overset{+}{N}H_3$

**3.** *Relationship between the Titration Curve and the Acid–Base Properties of Glycine* A 100 mL solution of 0.1 M glycine at pH 1.72 was titrated with 2 M NaOH solution. During the titration, the pH was monitored and the results were plotted in the graph shown. The key points in the titration are designated I to V on the graph. For each of the statements below, *identify* the appropriate key point in the titration and *justify* your choice.




(a) At what point will glycine be present predominantly as the species $^{+}H_3N—CH_2—COOH$?

(b) At what point is the *average* net charge of glycine $+\frac{1}{2}$?

(c) At what point is the amino group of half of the molecules ionized?

(d) At what point is the pH equal to the $pK_a$ of the carboxyl group?

(e) At what point is the pH equal to the $pK_a$ of the protonated amino group?

(f) At what points does glycine have its maximum buffering capacity?

(g) At what point is the *average* net charge zero?

(h) At what point has the carboxyl group been completely titrated (first equivalence point)?

(i) At what point are half of the carboxyl groups ionized?

(j) At what point is glycine completely titrated (second equivalence point)?

(k) At what point is the structure of the predominant species $^{+}H_3N—CH_2—COO^{-}$?

(l) At what point do the structures of the predominant species correspond to a 50:50 mixture of $^{+}H_3N—CH_2—COO^{-}$ and $H_2N—CH_2—COO^{-}$?

(m) At what point is the *average* net charge of glycine −1?

(n) At what point do the structures of the predominant species consist of a 50:50 mixture of $^{+}H_3N—CH_2—COOH$ and $^{+}H_3N—CH_2—COO^{-}$?

(o) What point corresponds to the isoelectric point?

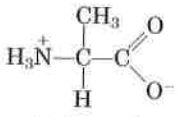
(p) At what point is the *average* net charge on glycine $-\frac{1}{2}$?

(q) What point represents the end of the titration?

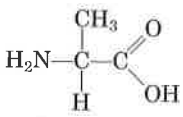
(r) If one wanted to use glycine as an efficient buffer, which points would represent the *worst* pH regions for buffering power?

(s) At what point in the titration is the predominant species $H_2N—CH_2—COO^{-}$?

**4.** *How Much Alanine Is Present as the Completely Uncharged Species?* At a pH equal to the isoelectric point, the *net* charge on alanine is zero. Two structures can be drawn that have a net charge of zero (zwitterionic and uncharged forms), but the predominant form of alanine at its pI is zwitterionic.



Zwitterionic



Uncharged

(a) Explain why the form of alanine at its pI is zwitterionic rather than completely uncharged.

(b) Estimate the fraction of alanine present at its pI as the completely uncharged form. Justify your assumptions.

**5.** *Ionization State of Amino Acids* Each ionizable group of an amino acid can exist in one of two states, charged or neutral. The electric charge on the functional group is determined by the relationship between its $pK_a$ and the pH of the solution. This relationship is described by the Henderson–Hasselbalch equation.

(a) Histidine has three ionizable functional groups. Write the relevant equilibrium equations for its three ionizations and assign the proper $pK_a$ for each ionization. Draw the structure of histidine in each ionization state. What is the net charge on the histidine molecule in each ionization state?

(b) Draw the structures of the predominant ionization state of histidine at pH 1, 4, 8, and 12. Note that the ionization state can be approximated by treating each ionizable group independently.

(c) What is the net charge of histidine at pH 1, 4, 8, and 12? For each pH, will histidine migrate toward the anode (+) or cathode (−) when placed in an electric field?

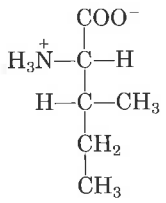
**6.** *Preparation of a Glycine Buffer* Glycine is commonly used as a buffer. Preparation of a 0.1 M glycine buffer starts with 0.1 M solutions of glycine hydrochloride ($HOOC—CH_2—NH_3^{+}Cl^{-}$) and gly-

cine ($^-OOC—CH_2—NH_3^+$), two commercially available forms of glycine. What volumes of these two solutions must be mixed to prepare 1 L of 0.1 M glycine buffer having a pH of 3.2? (Hint: See Box 4–2)

**7.** *Separation of Amino Acids by Ion-Exchange Chromatography* Mixtures of amino acids are analyzed by first separating the mixture into its components through ion-exchange chromatography. On a cation-exchange resin containing sulfonate groups (see Fig. 5–12), the amino acids flow down the column at different rates because of two factors that retard their movement: (1) ionic attraction between the $—SO_3^-$ residues on the column and positively charged functional groups on the amino acids and (2) hydrophobic interaction between amino acid side chains and the strongly hydrophobic backbone of the polystyrene resin. For each pair of amino acids listed, determine which member will be eluted first from an ion-exchange column by a pH 7.0 buffer.

(a) Asp and Lys
(b) Arg and Met
(c) Glu and Val
(d) Gly and Leu
(e) Ser and Ala

**8.** *Naming the Stereoisomers of Isoleucine* The structure of the amino acid isoleucine is:



(a) How many chiral centers does it have?
(b) How many optical isomers?
(c) Draw perspective formulas for all the optical isomers of isoleucine.

**9.** *Comparison of the $pK_a$ Values of an Amino Acid and Its Peptides* The titration curve of the amino acid alanine shows the ionization of two functional groups with $pK_a$ values of 2.34 and 9.69, corresponding to the ionization of the carboxyl and the protonated amino groups, respectively. The titration of di-, tri-, and larger oligopeptides of alanine also shows the ionization of only two functional groups, although the experimental $pK_a$ values are different. The trend in $pK_a$ values is summarized in the table.

| Amino acid or peptide | $pK_1$ | $pK_2$ |
|---|---|---|
| Ala | 2.34 | 9.69 |
| Ala–Ala | 3.12 | 8.30 |
| Ala–Ala–Ala | 3.39 | 8.03 |
| Ala–$(Ala)_n$–Ala, $n \geq 4$ | 3.42 | 7.94 |

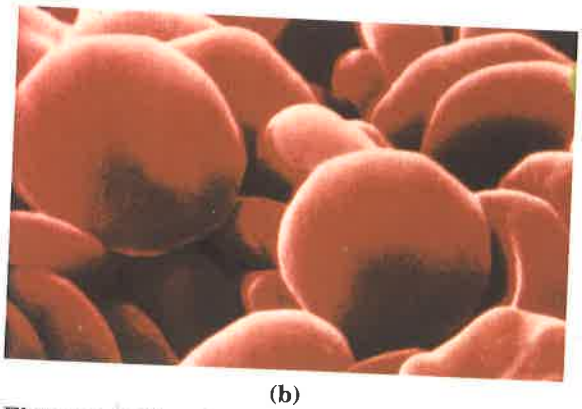
(a) Draw the structure of Ala–Ala–Ala. Identify the functional groups associated with $pK_1$ and $pK_2$.
(b) The value of $pK_1$ *increases* in going from Ala to an Ala oligopeptide. Provide an explanation for this trend.
(c) The value of $pK_2$ *decreases* in going from Ala to an Ala oligopeptide. Provide an explanation for this trend.

**10.** *Peptide Synthesis* In the synthesis of polypeptides on solid supports, the $\alpha$-amino group of each new amino acid is "protected" by a *t*-butyloxycarbonyl group (see Box 5–2). What would happen if this protecting group were not present?

CHAPTER

# 6 An Introduction to Proteins

(a)



(b)

**Figure 6–1** Functions of proteins. **(a)** The light produced by fireflies is the result of a light-producing reaction involving luciferin and ATP that is catalyzed by the enzyme luciferase (see Box 13–3). **(b)** Erythrocytes contain large amounts of the oxygen-transporting protein hemoglobin. **(c)** The white color of milk is derived primarily from the protein casein. **(d)** The movement of cilia in protozoans depends on the action of the protein dynein. **(e)** The protein fibroin is the major structural component of spider webs. **(f)** Castor beans contain a highly toxic protein called ricin. **(g)** Cancerous tumors are often made up of cells that have defects involving one or more of the proteins that regulate cell division.

Almost everything that occurs in the cell involves one or more proteins. Proteins provide structure, catalyze cellular reactions, and carry out a myriad of other tasks. Their central place in the cell is reflected in the fact that genetic information is ultimately expressed as protein. For each protein there is a segment of DNA (a gene; see Chapters 12 and 23) that encodes information specifying its sequence of amino acids. There are thousands of different kinds of proteins in a typical cell, each encoded by a gene and each performing a specific function. Proteins are among the most abundant biological macromolecules and are also extremely versatile in their functions.

The chapter begins with a discussion of some of the general properties of proteins. This is followed by a short summary of some common techniques used to purify and study proteins. Finally, we will examine the **primary structure** of protein molecules: the covalent backbone structure and the sequence of amino acid residues. One goal is to discover the relationships between amino acid sequence and biological function.

## Properties of Proteins

An understanding of these important macromolecules must begin with the fundamentals. What do proteins do? How big are they? What forms or shapes do they take? What are their chemical properties? The answers serve as an orientation to much that follows.

### Proteins Have Many Different Biological Functions

We can classify proteins according to their biological roles.

***Enzymes*** The most varied and most highly specialized proteins are those with catalytic activity—the enzymes. Virtually all the chemical reactions of organic biomolecules in cells are catalyzed by enzymes. Many thousands of different enzymes, each capable of catalyzing a different kind of chemical reaction, have been discovered in different organisms (Fig. 6–1a).

***Transport Proteins*** Transport proteins in blood plasma bind and carry specific molecules or ions from one organ to another. Hemoglobin of erythrocytes (Fig. 6–1b) binds oxygen as the blood passes through the lungs, carries it to the peripheral tissues, and there releases it to participate in the energy-yielding oxidation of nutrients. The blood

plasma contains lipoproteins, which carry lipids from the liver to other organs. Other kinds of transport proteins are present in the plasma membranes and intracellular membranes of all organisms; these are adapted to bind glucose, amino acids, or other substances and transport them across the membrane.

***Nutrient and Storage Proteins*** The seeds of many plants store nutrient proteins required for the growth of the germinating seedling. Particularly well-studied examples are the seed proteins of wheat, corn, and rice. Ovalbumin, the major protein of egg white, and casein, the major protein of milk, are other examples of nutrient proteins (Fig. 6–1c). The ferritin found in some bacteria and in plant and animal tissues stores iron.

***Contractile or Motile Proteins*** Some proteins endow cells and organisms with the ability to contract, to change shape, or to move about. Actin and myosin function in the contractile system of skeletal muscle and also in many nonmuscle cells. Tubulin is the protein from which microtubules are built. Microtubules act in concert with the protein dynein in flagella and cilia (Fig. 6–1d) to propel cells.

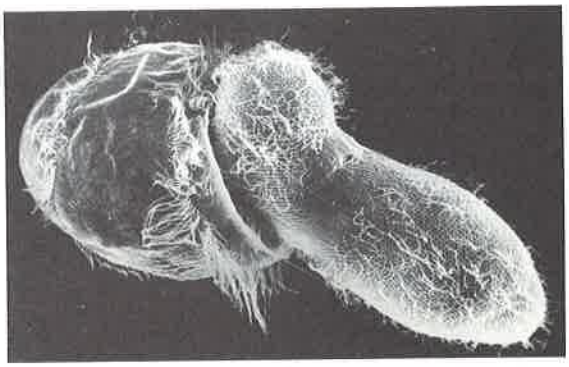
***Structural Proteins*** Many proteins serve as supporting filaments, cables, or sheets, to give biological structures strength or protection. The major component of tendons and cartilage is the fibrous protein collagen, which has very high tensile strength. Leather is almost pure collagen. Ligaments contain elastin, a structural protein capable of stretching in two dimensions. Hair, fingernails, and feathers consist largely of the tough, insoluble protein keratin. The major component of silk fibers and spider webs is fibroin (Fig. 6–1e). The wing hinges of some insects are made of resilin, which has nearly perfect elastic properties.

***Defense Proteins*** Many proteins defend organisms against invasion by other species or protect them from injury. The immunoglobulins or antibodies, specialized proteins made by the lymphocytes of vertebrates, can recognize and precipitate or neutralize invading bacteria, viruses, or foreign proteins from another species. Fibrinogen and thrombin are blood-clotting proteins that prevent loss of blood when the vascular system is injured. Snake venoms, bacterial toxins, and toxic plant proteins, such as ricin, also appear to have defensive functions (Fig. 6–1f). Some of these, including fibrinogen, thrombin, and some venoms, are also enzymes.

***Regulatory Proteins*** Some proteins help regulate cellular or physiological activity. Among them are many hormones. Examples include insulin, which regulates sugar metabolism, and the growth hormone of the pituitary. The cellular response to many hormonal signals is often mediated by a class of GTP-binding proteins called G proteins (GTP is closely related to ATP, with guanine replacing the adenine portion of the molecule; see Figs. 1–12 and 3–16b.) Other regulatory proteins bind to DNA and regulate the biosynthesis of enzymes and RNA molecules involved in cell division in both prokaryotes and eukaryotes (Fig. 6–1g).

***Other Proteins*** There are numerous other proteins whose functions are rather exotic and not easily classified. Monellin, a protein of an African plant, has an intensely sweet taste. It is being studied as a

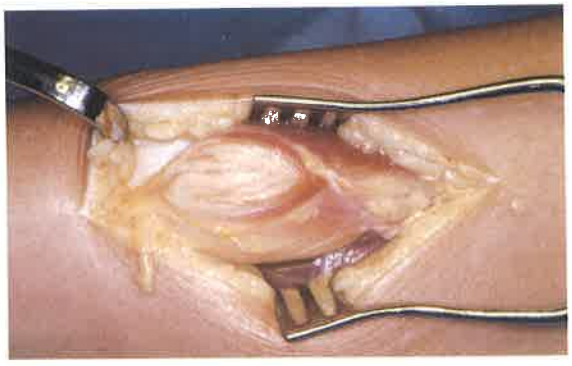
(c)



(d)

(e)

(f)



(g)

nonfattening, nontoxic food sweetener for human use. The blood plasma of some Antarctic fish contains antifreeze proteins, which protect their blood from freezing.

It is extraordinary that all these proteins, with their very different properties and functions, are made from the same group of 20 amino acids.

## Proteins Are Very Large Molecules

How long are the polypeptide chains in proteins? Table 6–1 shows that human cytochrome *c* has 104 amino acid residues linked in a single chain; bovine chymotrypsinogen has 245 amino acid residues. Probably near the upper limit of size is the protein apolipoprotein B, a cholesterol-transport protein with 4,536 amino acid residues in a single polypeptide chain of molecular weight 513,000. Most naturally occurring polypeptides contain less than 2,000 amino acid residues.

**Table 6–1 Molecular data on some proteins**

| | Molecular weight | Number of residues | Number of polypeptide chains |
|---|---|---|---|
| Insulin (bovine) | 5,733 | 51 | 2 |
| Cytochrome *c* (human) | 13,000 | 104 | 1 |
| Ribonuclease A (bovine pancreas) | 13,700 | 124 | 1 |
| Lysozyme (egg white) | 13,930 | 129 | 1 |
| Myoglobin (equine heart) | 16,890 | 153 | 1 |
| Chymotrypsin (bovine pancreas) | 21,600 | 241 | 3 |
| Chymotrypsinogen (bovine) | 22,000 | 245 | 1 |
| Hemoglobin (human) | 64,500 | 574 | 4 |
| Serum albumin (human) | 68,500 | ~550 | 1 |
| Hexokinase (yeast) | 102,000 | ~800 | 2 |
| Immunoglobulin G (human) | 145,000 | ~1,320 | 4 |
| RNA polymerase (*E. coli*) | 450,000 | ~4,100 | 5 |
| Apolipoprotein B (human) | 513,000 | 4,536 | 1 |
| Glutamate dehydrogenase (bovine liver) | 1,000,000 | ~8,300 | ~40 |

Some proteins consist of a single polypeptide chain, but others, called **multisubunit** proteins, have two or more (Table 6–1). The individual polypeptide chains in a multisubunit protein may be identical or different. If at least some are identical, the protein is sometimes called an **oligomeric** protein and the subunits themselves are referred to as **protomers.** The enzyme ribonuclease has one polypeptide chain. Hemoglobin has four: two identical $\alpha$ chains and two identical $\beta$ chains, all four held together by noncovalent interactions.

The molecular weights of proteins, which can be determined by various physicochemical methods, may range from little more than 10,000 for small proteins such as cytochrome *c* (104 residues), to more than $10^6$ for proteins with very long polypeptide chains or those with several subunits. The molecular weights of some typical proteins are given in Table 6–1. No simple generalizations can be made about the molecular weights of proteins in relation to their function.

One can calculate the approximate number of amino acid residues in a simple protein containing no other chemical group by dividing its molecular weight by 110. Although the average molecular weight of the

20 standard amino acids is about 138, the smaller amino acids predominate in most proteins; when weighted for the proportions in which the various amino acids occur in proteins (see Table 5–1), the average molecular weight is nearer to 128. Because a molecule of water ($M_r$ 18) is removed to create each peptide bond, the average molecular weight of an amino acid residue in a protein is about 128 − 18 = 110. Table 6–1 shows the number of amino acid residues in several proteins.

## Proteins Have Characteristic Amino Acid Compositions

As is true for simple peptides, hydrolysis of proteins with acid or base yields a mixture of free α-amino acids. When completely hydrolyzed, each type of protein yields a characteristic proportion or mixture of the different amino acids. Table 6–2 shows the composition of the amino acid mixtures obtained on complete hydrolysis of human cytochrome *c* and of bovine chymotrypsinogen, the inactive precursor of the digestive enzyme chymotrypsin. These two proteins, with very different functions, also differ significantly in the relative numbers of each kind of amino acid they contain. The 20 amino acids almost never occur in equal amounts in proteins. Some amino acids may occur only once per molecule or not at all in a given type of protein; others may occur in large numbers.

**Table 6–2 Amino acid composition of two proteins**

| Amino acid | Number of residues per molecule of protein: Human cytochrome *c* | Bovine chymotrypsinogen |
|---|---|---|
| Ala | 6 | 22 |
| Arg | 2 | 4 |
| Asn | 5 | 15 |
| Asp | 3 | 8 |
| Cys | 2 | 10 |
| Gln | 2 | 10 |
| Glu | 8 | 5 |
| Gly | 13 | 23 |
| His | 3 | 2 |
| Ile | 8 | 10 |
| Leu | 6 | 19 |
| Lys | 18 | 14 |
| Met | 3 | 2 |
| Phe | 3 | 6 |
| Pro | 4 | 9 |
| Ser | 2 | 28 |
| Thr | 7 | 23 |
| Trp | 1 | 8 |
| Tyr | 5 | 4 |
| Val | 3 | 23 |
| Total | 104 | 245 |

## Some Proteins Contain Chemical Groups Other Than Amino Acids

Many proteins, such as the enzymes ribonuclease and chymotrypsinogen, contain only amino acids and no other chemical groups; these are considered simple proteins. However, some proteins contain chemical components in addition to amino acids; these are called **conjugated proteins.** The non–amino acid part of a conjugated protein is usually called its **prosthetic group.** Conjugated proteins are classified on the basis of the chemical nature of their prosthetic groups (Table 6–3); for example, **lipoproteins** contain lipids, **glycoproteins** contain sugar groups, and **metalloproteins** contain a specific metal. A number of proteins contain more than one prosthetic group. Usually the prosthetic group plays an important role in the protein's biological function.

**Table 6–3 Conjugated proteins**

| Class | Prosthetic group | Example |
|---|---|---|
| Lipoproteins | Lipids | $\beta_1$-Lipoprotein of blood |
| Glycoproteins | Carbohydrates | Immunoglobulin G |
| Phosphoproteins | Phosphate groups | Casein of milk |
| Hemoproteins | Heme (iron porphyrin) | Hemoglobin |
| Flavoproteins | Flavin nucleotides | Succinate dehydrogenase |
| Metalloproteins | Iron | Ferritin |
| | Zinc | Alcohol dehydrogenase |
| | Calcium | Calmodulin |
| | Molybdenum | Dinitrogenase |
| | Copper | Plastocyanin |

## Working with Proteins

The aggregate biochemical picture of protein structure and function is derived from the study of many individual proteins. To study a protein in any detail it must be separated from all other proteins in a cell, and techniques must be available to determine its properties. The necessary methods come from protein chemistry, a discipline as old as biochemistry itself and one that retains a central position in biochemical research. Modern techniques are providing ever newer experimental insights into the critical relationship between the structure of a protein and its function.

### Proteins Can Be Separated and Purified

Cells contain thousands of different kinds of proteins. A pure preparation of a given protein is essential before its properties, amino acid composition, and sequence can be determined. How, then, can one protein be purified?

Methods for separating proteins take advantage of properties such as charge, size, and solubility, which vary from one protein to the next. Because many proteins bind to other biomolecules, proteins can also be separated on the basis of their binding properties. The source of a protein is generally tissue or microbial cells. The cells must be broken open and the protein released into a solution called a **crude extract.** If necessary, differential centrifugation can be used to prepare subcellular fractions or to isolate organelles (see Fig. 2–24). Once the extract or organelle preparation is ready, a variety of methods are available for separation of proteins. Ion-exchange chromatography (see Fig. 5–12) can be used to separate proteins with different charges in much the same way that it separates amino acids. Other chromatographic methods take advantage of differences in size, binding affinity, and solubility (Fig. 6–2). Nonchromatographic methods include the selective precipitation of proteins with salt, acid, or high temperatures.

The approach to the purification of a "new" protein, one not previously isolated, is guided both by established precedents and common sense. In most cases, several different methods must be used sequentially to completely purify a protein. The choice of method is somewhat empirical, and many protocols may be tried before the most effective is determined. Trial and error can often be minimized by using purification procedures developed for similar proteins as a guide. Published purification protocols are available for many thousands of proteins. Common sense dictates that inexpensive procedures be used first, when the total volume and number of contaminants is greatest. Chromatographic methods are often impractical at early stages because the amount of chromatographic medium needed increases with sample size. As each purification step is completed, the sample size generally becomes smaller (Table 6–4) and more sophisticated (and expensive) chromatographic procedures can be applied.

### Individual Proteins Can Be Quantified

In order to purify a protein, it is essential to have an assay to detect and quantify that protein in the presence of many other proteins. Often, purification must proceed in the absence of any information about the size and physical properties of the protein, or the fraction of the total protein mass it represents in the extract.

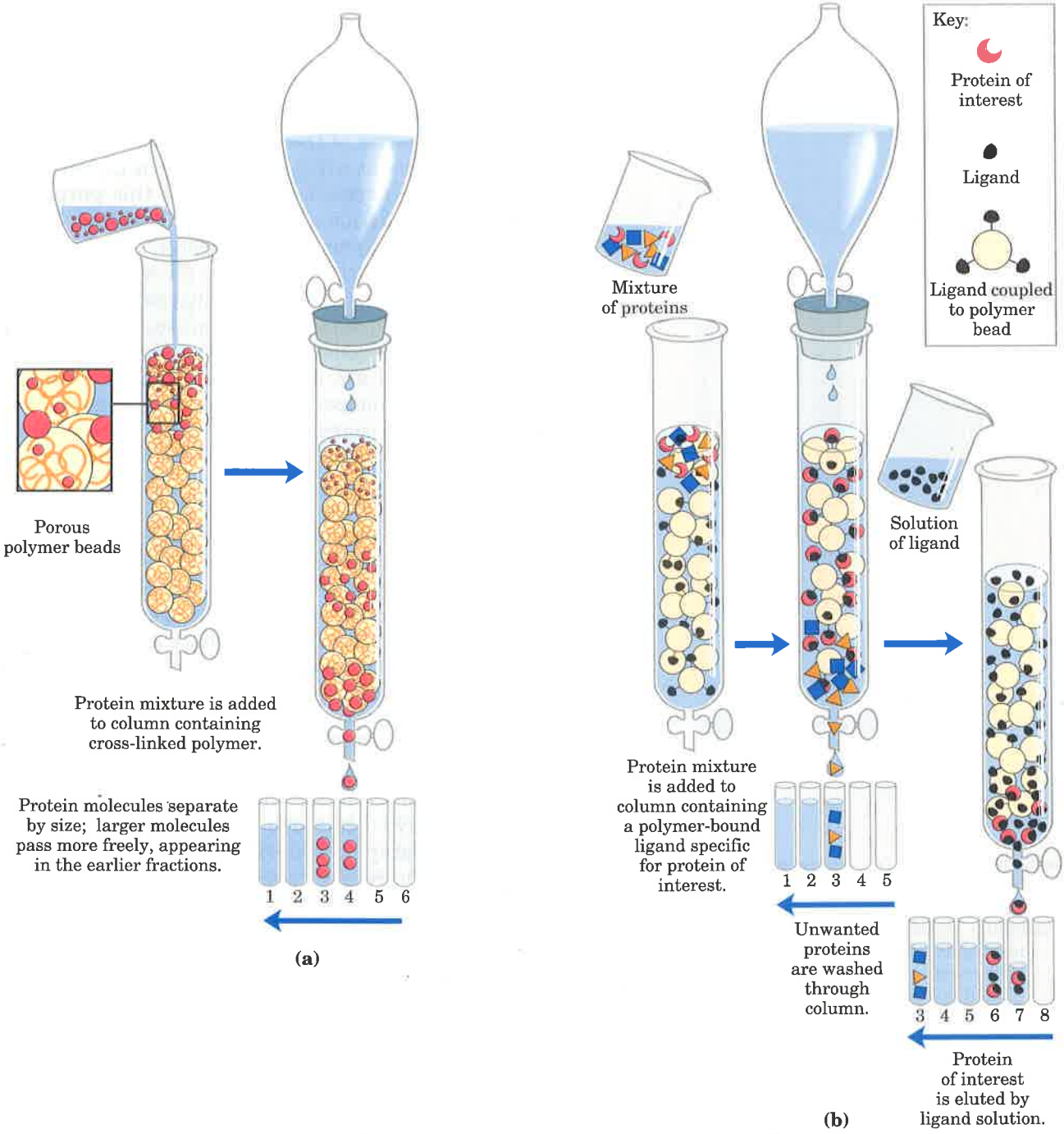



**Table 6–4 A purification table for a hypothetical enzyme***

| Procedure or step | Fraction volume (ml) | Total protein (mg) | Activity (units) | Specific activity (units/mg) |
|---|---|---|---|---|
| 1. Crude cellular extract | 1,400 | 10,000 | 100,000 | 10 |
| 2. Precipitation | 280 | 3,000 | 96,000 | 32 |
| 3. Ion-exchange chromatography | 90 | 400 | 80,000 | 200 |
| 4. Size-exclusion chromatography | 80 | 100 | 60,000 | 600 |
| 5. Affinity chromatography | 6 | 3 | 45,000 | 15,000 |

* All data represent the status of the sample *after* the procedure indicated in the first column has been carried out.

**Figure 6–2** Two types of chromatographic methods used in protein purification. **(a)** Size-exclusion chromatography; also called gel filtration. This method separates proteins according to size. The column contains a cross-linked polymer with pores of selected size. Larger proteins migrate faster than smaller ones, because they are too large to enter the pores in the beads and hence take a more direct route through the column. The smaller proteins enter the pores and are slowed by the more labyrinthian path they take through the column. **(b)** Affinity chromatography separates proteins by their binding specificities. The proteins retained on the column are those that bind specifically to a ligand cross-linked to the beads. (In biochemistry, the term "ligand" is used to refer to a group or molecule that is bound.) After nonspecific proteins are washed through the column, the bound protein of particular interest is eluted by a solution containing free ligand.

The amount of an enzyme in a given solution or tissue extract can be assayed in terms of the catalytic effect it produces, that is, the *increase* in the rate at which its substrate is converted to reaction products when the enzyme is present. For this purpose one must know (1) the overall equation of the reaction catalyzed, (2) an analytical procedure for determining the disappearance of the substrate or the appearance of the reaction products, (3) whether the enzyme requires cofactors such as metal ions or coenzymes, (4) the dependence of the enzyme activity on substrate concentration, (5) the optimum pH, and (6) a temperature zone in which the enzyme is stable and has high activity. Enzymes are usually assayed at their optimum pH and at some convenient temperature within the range 25 to 38 °C. Also, very high substrate concentrations are generally required so that the initial reaction rate, which is measured experimentally, is proportional to enzyme concentration (Chapter 8).

By international agreement, 1.0 unit of enzyme activity is defined as the amount of enzyme causing transformation of 1.0 μmol of substrate per minute at 25 °C under optimal conditions of measurement. The term **activity** refers to the total units of enzyme in the solution. The **specific activity** is the number of enzyme units per milligram of protein (Fig. 6–3). The specific activity is a measure of enzyme purity: it increases during purification of an enzyme and becomes maximal and constant when the enzyme is pure (Table 6–4).

After each purification step, the activity of the preparation (in units) is assayed, the total amount of protein is determined independently, and their ratio gives the specific activity. Activity and total protein generally decrease with each step. Activity decreases because some loss always occurs due to inactivation or nonideal interactions with chromatographic materials or other molecules in the solution. Total protein decreases because the objective is to remove as much nonspecific protein as possible. In a successful step, the loss of nonspecific protein is much greater than the loss of activity; therefore, specific activity increases even as total activity falls. The data are then assembled in a purification table (Table 6–4). A protein is generally considered pure when further purification steps fail to increase specific activity, and when only a single protein species can be detected (by methods to be described later).

For proteins that are not enzymes, other quantification methods are required. Transport proteins can be assayed by their binding to the molecule they transport, and hormones and toxins by the biological effect they produce; for example, growth hormones will stimulate the growth of certain cultured cells. Some structural proteins represent such a large fraction of a tissue mass that they can be readily extracted and purified without an assay. The approaches are as varied as the proteins themselves.

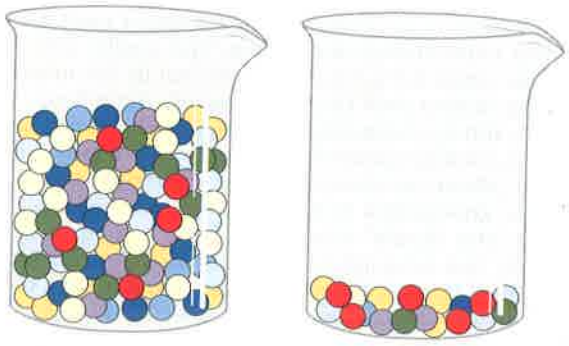
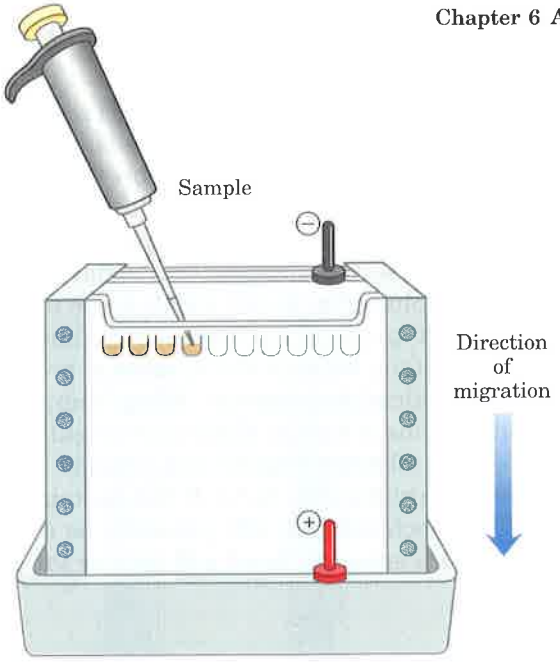
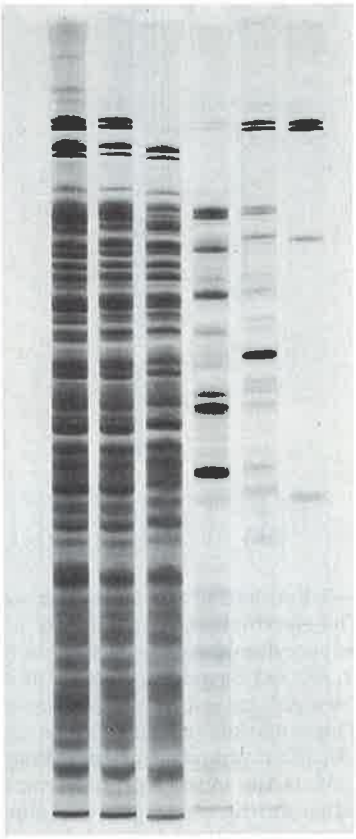


**Figure 6–3** Activity versus specific activity. The difference between these two terms can be illustrated by considering two jars of marbles. The jars contain the same number of red marbles (representing an unknown protein), but different amounts of marbles of other colors. If the marbles are taken to represent proteins, both jars contain the same *activity* of the protein represented by the red marbles. The second jar, however, has the higher *specific activity* because here the red marbles represent a much higher fraction of the total.




(a)



(b)

**Figure 6–4** Electrophoresis. **(a)** Different samples are loaded in wells or depressions at the top of the polyacrylamide gel. The proteins move into the gel when an electric field is applied. The gel minimizes convection currents caused by small temperature gradients, and it minimizes protein movements other than those induced by the electric field. **(b)** Proteins can be visualized after electrophoresis by treating the gel with a stain such as Coomassie blue, which binds to the proteins but not to the gel itself. Each band on the gel represents a different protein (or protein subunit); smaller proteins are found nearer the bottom of the gel. This gel illustrates the purification of the enzyme RNA polymerase from the bacterium *E. coli*. The first lane shows the proteins present in the crude cellular extract. Successive lanes show the proteins present after each purification step. The purified protein contains four subunits, as seen in the last lane on the right.

## Proteins Can Be Characterized by Electrophoresis

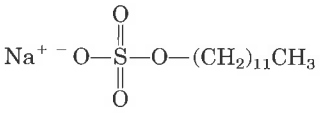
In addition to chromatography, another important set of methods is available for the separation of proteins, based on the migration of charged proteins in an electric field, a process called **electrophoresis.** These procedures are not often used to purify proteins in large amounts because simpler alternative methods are usually available and electrophoretic methods often inactivate proteins. Electrophoresis is, however, especially useful as an analytical method. Its advantage is that proteins can be visualized as well as separated, permitting a researcher to estimate quickly the number of proteins in a mixture or the degree of purity of a particular protein preparation. Also, electrophoresis allows determination of crucial properties of a protein such as its isoelectric point and approximate molecular weight.

In electrophoresis, the force moving the macromolecule (nucleic acids as well as proteins are separated this way) is the electrical potential, $E$. The electrophoretic mobility of the molecule, $\mu$, is the ratio of the velocity of the particle, $V$, to the electrical potential. Electrophoretic mobility is also equal to the net charge of the molecule, $Z$, divided by the frictional coefficient, $f$. Thus:

$$\mu = \frac{V}{E} = \frac{Z}{f}$$

Electrophoresis of proteins is generally carried out in gels made up of the cross-linked polymer polyacrylamide (Fig. 6–4). The polyacrylamide gel acts as a molecular sieve, slowing the migration of proteins approximately in proportion to their mass, or molecular weight.

An electrophoretic method commonly used for estimation of purity and molecular weight makes use of the detergent **sodium dodecyl sulfate** (SDS). SDS binds to most proteins (probably by hydrophobic interactions; see Chapter 4) in amounts roughly proportional to the molecular weight of the protein, about one molecule of SDS for every two amino acid residues. The bound SDS contributes a large net negative charge, rendering the intrinsic charge of the protein insignificant.




Sodium dodecyl sulfate (SDS)

In addition, the native conformation of a protein is altered when SDS is bound, and most proteins assume a similar shape, and thus a similar ratio of charge to mass. Electrophoresis in the presence of SDS therefore separates proteins almost exclusively on the basis of mass (molecular weight), with smaller polypeptides migrating more rapidly. After electrophoresis, the proteins are visualized by adding a dye such as Coomassie blue (Fig. 6–4b) which binds to proteins but not to the gel itself. This type of gel provides one method to monitor progress in isolating a protein, because the number of protein bands should decrease as the purification proceeds. When compared with the positions to which proteins of known molecular weight migrate in the gel, the position of an unknown protein can provide an excellent measure of its molecular weight (Fig. 6–5). If the protein has two or more different subunits, each subunit will generally be separated by the SDS treatment, and a separate band will appear for each.

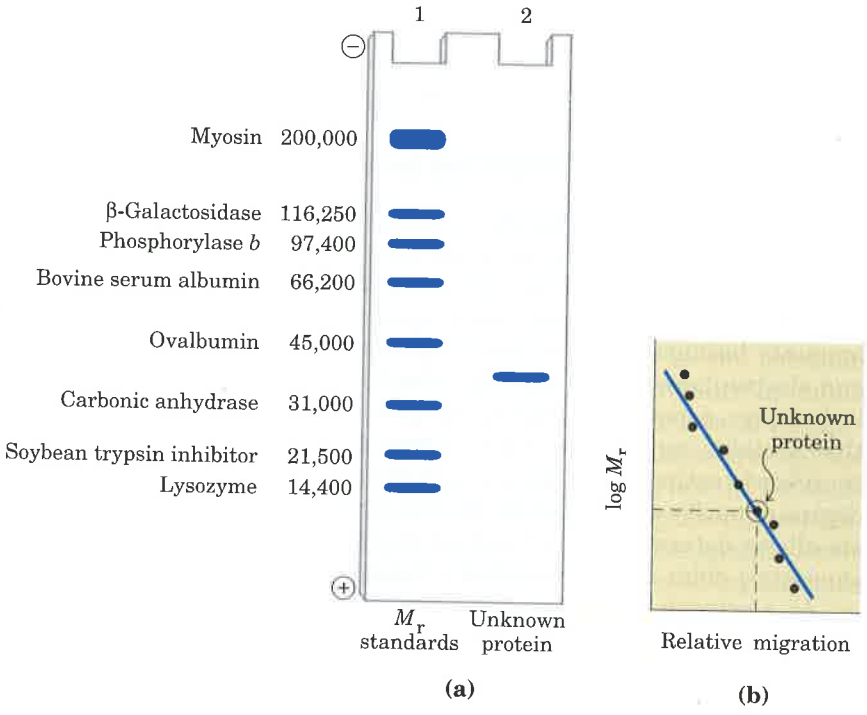



**Figure 6–5** Estimating the molecular weight of a protein. The electrophoretic mobility of a protein on an SDS polyacrylamide gel is related to its molecular weight, $M_r$. **(a)** Standard proteins of known molecular weight are subjected to electrophoresis (lane 1). These marker proteins can be used to estimate the $M_r$ of an unknown protein (lane 2). **(b)** A plot of log $M_r$ of the marker proteins versus relative migration during electrophoresis allows the $M_r$ of the unknown protein to be read from the graph.

**Table 6–5 The isoelectric points of some proteins**

| | pI |
|---|---|
| Pepsin | ~1.0 |
| Egg albumin | 4.6 |
| Serum albumin | 4.9 |
| Urease | 5.0 |
| β-Lactoglobulin | 5.2 |
| Hemoglobin | 6.8 |
| Myoglobin | 7.0 |
| Chymotrypsinogen | 9.5 |
| Cytochrome *c* | 10.7 |
| Lysozyme | 11.0 |

**Isoelectric focusing** is a procedure used to determine the isoelectric point (pI) of a protein (Fig. 6–6). A pH gradient is established by allowing a mixture of low molecular weight organic acids and bases (ampholytes; see p. 118) to distribute themselves in an electric field generated across the gel. When a protein mixture is applied, each protein migrates until it reaches the pH that matches its pI. Proteins with different isoelectric points are thus distributed differently throughout the gel (Table 6–5).

Combining these two electrophoretic methods in two-dimensional gels permits the resolution of complex mixtures of proteins (Fig. 6–7). This is a more sensitive analytical method than either isoelectric focusing or SDS electrophoresis alone. Two-dimensional electrophoresis separates proteins of identical molecular weight that differ in pI, or proteins with similar pI values but different molecular weights.

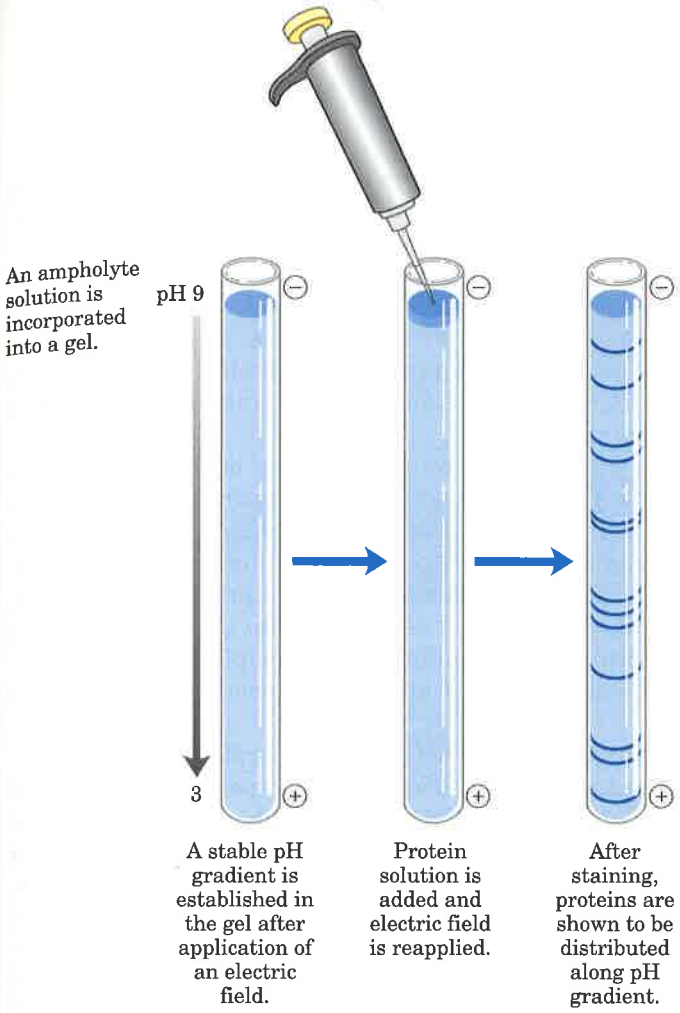
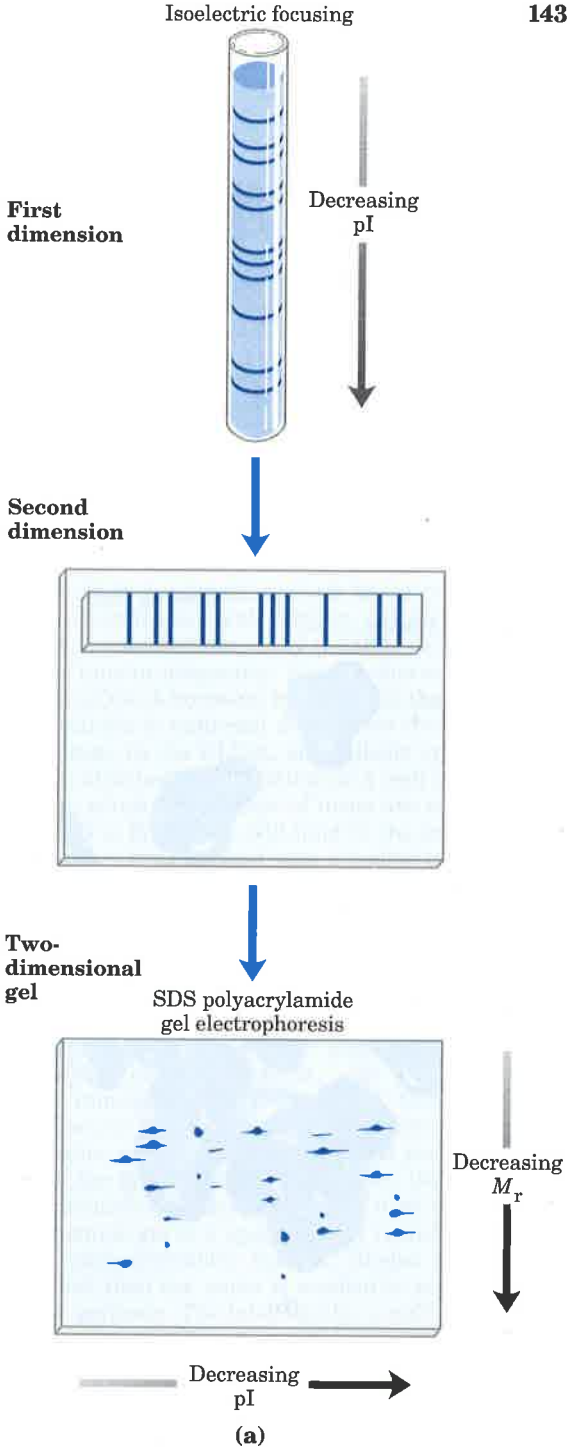



**Figure 6–6** Isoelectric focusing. This technique separates proteins according to their isoelectric points. A stable pH gradient is established in the gel by the addition of appropriate ampholytes. A protein mixture is placed in a well on the gel. With an applied electric field, proteins enter the gel and migrate until each reaches a pH equivalent to its pI. Remember that the net charge of a protein is zero when pH = pI.



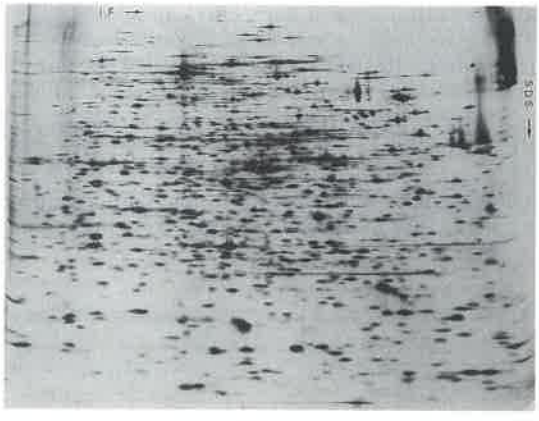
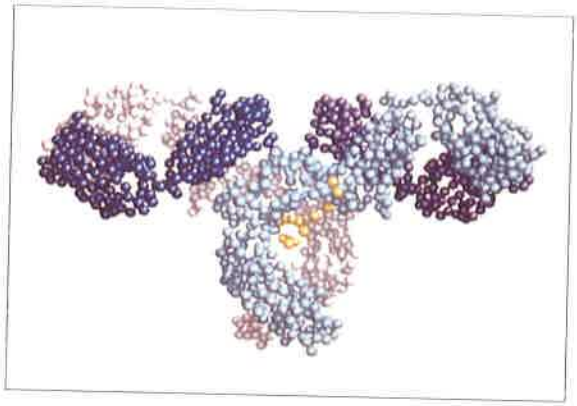
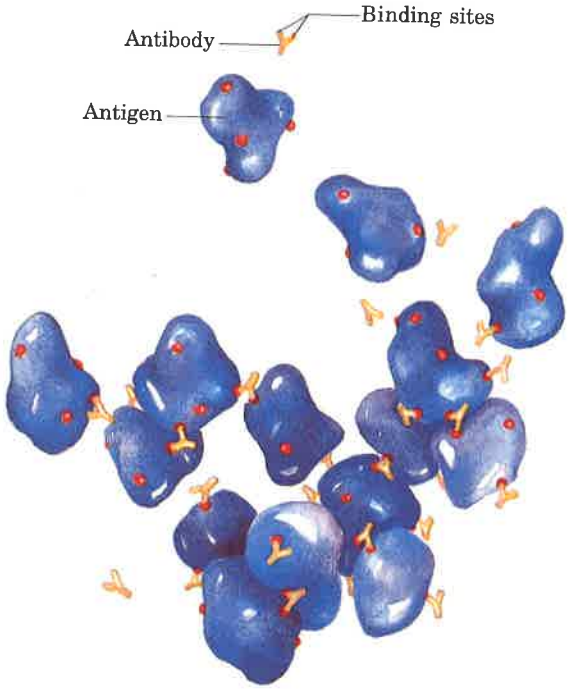




**Figure 6–7** Two-dimensional electrophoresis. **(a)** Proteins are first separated by isoelectric focusing. The gel is then laid horizontally on a second gel, and the proteins are separated by SDS polyacrylamide gel electrophoresis. In this two-dimensional gel, horizontal separation reflects differences in pI; vertical separation reflects differences in molecular weight. **(b)** More than 1,000 different proteins from *E. coli* can be resolved using this technique.



(a)




(b)

**Figure 6–8** The immune response and the action of antibodies. **(a)** A molecule of immunoglobulin G (IgG) consists of two polypeptides known as heavy chains (white and light blue) and two known as light chains (purple and dark blue). Immunoglobulins are glycoproteins and contain bound carbohydrate (yellow). **(b)** Each antigen evokes a specific set of antibodies, which will recognize and combine only with that antigen or closely related molecules. (Antibody-binding sites are shown as red areas on the antigen.) The Y-shaped antibodies each have two binding sites for the antigen, and can precipitate the antigen by forming an insoluble, latticelike aggregate.

## The Antibody–Antigen Interaction Is Used to Quantify and Localize Proteins

Several sensitive analytical procedures have been developed from the study of a class of proteins called **antibodies** or **immunoglobulins.** Antibody molecules appear in the blood serum and certain tissues of a vertebrate animal in response to injection of an **antigen,** a protein or other macromolecule foreign to that individual. Each foreign protein elicits the formation of a set of different antibodies, which can combine with the antigen to form an antigen–antibody complex. The production of antibodies is part of a general defense mechanism in vertebrates called the **immune response.**

Antibodies are Y-shaped proteins consisting of four polypeptide chains. They have two binding sites that are complementary to specific structural features of the antigen molecule, making possible the formation of a three-dimensional lattice of alternating antigen and antibody molecules (Fig. 6–8). If sufficient antigen is present in a sample, the addition of antibodies or blood serum from an immunized animal will result in the formation of a quantifiable precipitate. No such precipitate is formed when serum of an unimmunized animal is mixed with the antigen.

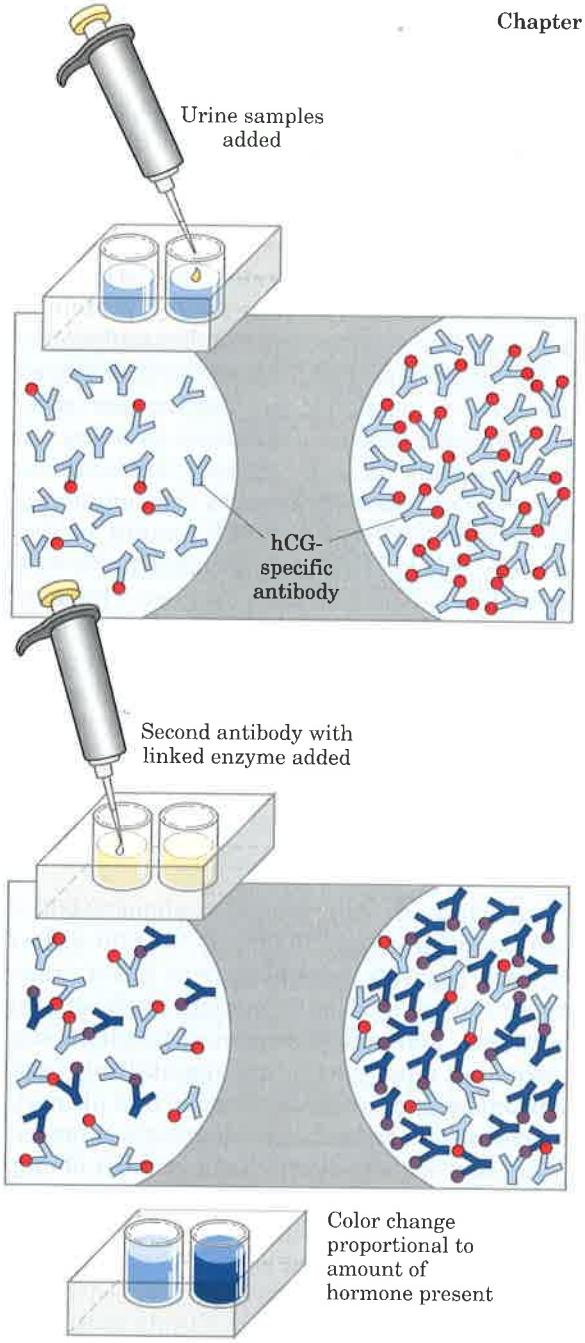
Antibodies are highly specific for the foreign proteins or other macromolecules that evoke their formation. It is this specificity that makes them valuable analytical reagents. A rabbit antibody formed to horse serum albumin, for example, will combine with the latter but will not usually combine with other horse proteins, such as horse hemoglobin.

Two types of antibody preparations are in use: **polyclonal** and **monoclonal.** Polyclonal antibodies are those produced by many different types (or populations) of antibody-producing cells in an animal immunized with an antigen (in this case a protein). Each type of cell produces an antibody that binds only to a specific, small part of the antigen protein. Consequently, polyclonal preparations contain a mixture of antibodies that recognize different parts of the protein. Monoclonal antibodies, in contrast, are synthesized by a population of identical cells (a **clone**) grown in cell culture. These antibodies are homogeneous, all recognizing the same specific part of the protein. The techniques for producing monoclonal antibodies were worked out by Georges Köhler and Cesar Milstein.

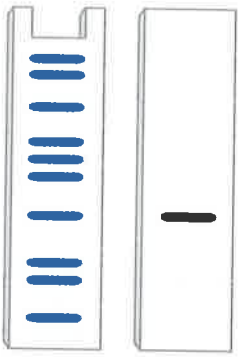
Antibodies are so exquisitely specific that they can in some cases distinguish between two proteins differing by only a single amino acid.

Cesar Milstein

Georges Köhler




(a)




(b)

(c)

**Figure 6–9** Analytical methods based on the interaction of antibodies with antigen. **(a)** An enzyme-linked immunosorbent assay (ELISA) used in testing for human pregnancy. Human chorionic gonadotropin (hCG), a hormone produced by the placenta, is detectable in maternal urine a few days after conception. In the ELISA, an antibody specific for hCG is attached to the bottom of a well in a plastic tray, to which a few drops of urine are added. If any hCG is present, it will bind to the antibodies. The tube is then washed, and a second antibody (also specific for hCG) is added. This second antibody is linked to an enzyme that catalyzes the conversion of a colorless compound to a colored one; the amount of colored compound produced provides a sensitive measure of the amount of hormone present. The ELISA has been adapted for use in determining the amount of specific proteins in tissue samples, in blood, or in urine.

**(b)** Immunoblot (or Western blot) technique. Proteins are separated by electrophoresis, then antibodies are used to determine the presence and size of the proteins. After separation, the proteins are transferred electrophoretically from an SDS polyacrylamide gel to a special paper (which makes them more accessible). Specific, labeled antibody is added, then the paper is washed to remove unbound antibody. The label can be a radioactive element, a fluorescent compound, or an enzyme as in the ELISA. The position of the labeled antibody defines the $M_r$ of the detected protein. All of the proteins are seen in the stained gel; only the protein bound to the antibody is seen in the immunoblot.

**(c)** In immunocytochemistry, labeled antibodies are introduced into cells to reveal the subcellular location of a specific protein. Here, fluorescently labeled antibodies and a fluorescence microscope have been used to locate tubulin filaments in a human fibroblast.

When a mixture of proteins is added to a chromatography column in which the antibody is covalently attached to a resin, the antibody will specifically bind its target protein and retain it on the column while other proteins are washed through. The target protein can then be eluted from the resin by a salt solution or some other agent. This can be a powerful tool for protein purification.

A variety of other analytical techniques rely on antibodies. In each case the antibody is attached to a radioactive label or some other reagent to make it easy to detect. The antibody binds the target protein, and the label reveals its presence in a solution or its location in a gel or even a living cell. Several variations of this procedure are illustrated in Figure 6–9. We shall examine some other aspects of antibodies in chapters to follow; they are of extreme importance in medicine and also tell much about the structure of proteins and the action of genes.

## The Covalent Structure of Proteins

All proteins in all species, regardless of their function or biological activity, are built from the same set of 20 amino acids (Chapter 5). What is it, then, that makes one protein an enzyme, another a hormone, another a structural protein, and still another an antibody? How do they differ chemically? Quite simply, proteins differ from each other because each has a distinctive number and *sequence* of amino acid residues. The amino acids are the alphabet of protein structure; they can be arranged in an almost infinite number of sequences to make an almost infinite number of different proteins. A specific sequence of amino acids folds up into a unique three-dimensional structure, and this structure in turn determines the function of the protein.

The amino acid sequence of a protein, or its **primary structure,** can be very informative to a biochemist. No other property so clearly distinguishes one protein from another. This now becomes the focus of the remainder of the chapter. We first consider empirical clues that amino acid sequence and protein function are closely linked, then describe how amino acid sequence is determined, and finally outline the many uses to which this information can be put.

### The Function of a Protein Depends on Its Amino Acid Sequence

The bacterium *E. coli* produces about 3,000 different proteins. A human being produces 50,000 to 100,000 different proteins. In both cases, each separate type of protein has a unique structure and this structure confers a unique function. Each separate type of protein also has a unique amino acid sequence. Intuition suggests that the amino acid sequence must play a fundamental role in determining the three-dimensional structure of the protein, and ultimately its function, but is this expectation correct? A quick survey of proteins and how they vary in amino acid sequence provides a number of empirical clues that help substantiate the important relationship between amino acid sequence and biological function. First, as we have already noted, proteins with different functions always have different amino acid sequences. Second, more than 1,400 human genetic diseases have been traced to the production of defective proteins (Table 6–6). Perhaps a third of these proteins are defective because of a single change in the amino acid sequence; hence, if the primary structure is altered, the function of the protein may also be changed. Finally, on comparing proteins with similar functions from different species, we find that these proteins often have similar amino acid sequences. An extreme case is ubiquitin, a 76 amino acid protein involved in regulating the degradation of other proteins. The amino acid sequence of ubiquitin is identical in species as disparate as fruit flies and humans.

Is the amino acid sequence absolutely fixed, or invariant, for a particular protein? No; some flexibility is possible. An estimated 20 to 30% of the proteins in humans are **polymorphic,** having amino acid sequence variants in the human population. Many of these variations in sequence have little or no effect on the function of the protein. Furthermore, proteins that carry out a broadly similar function in distantly related species often differ greatly in overall size and amino acid

sequence. An example is DNA polymerase, the primary enzyme involved in DNA synthesis. The DNA polymerase of a bacterium is very different in much of its sequence from that of a mouse cell.

**Table 6–6 A sampling of genetic diseases linked to loss or defect of a single enzyme or protein**

| Disease | Physiological effects | Affected enzyme or protein |
|---|---|---|
| Cystic fibrosis | Abnormal secretion in lungs, pancreas, sweat glands; chronic pulmonary disease generally leading to death in children or young adults | Chloride channel |
| Lesch–Nyhan syndrome | Neurological defects, self-mutilation, mental retardation | Hypoxanthine-guanine phosphoribosyl transferase |
| Immunodeficiency disease | Severe loss of immune response | Purine nucleoside phosphorylase |
| Immunodeficiency disease | Severe loss of immune response (children must live in a sterile bubble) | Adenosine deaminase |
| Gaucher's disease | Erosion of bones, hip joints; sometimes brain damage | Glucocerebrosidase |
| Gout, primary | Overproduction of uric acid resulting in recurring attacks of acute arthritis | Phosphoribosyl pyrophosphate synthetase |
| Rickets, vitamin D-dependent | Short stature, convulsions | 25-Hydroxycholecalciferol-1-hydroxylase |
| Familial hypercholesterolemia | Atherosclerosis resulting from elevated cholesterol levels in blood; sometimes early death from heart failure | Low-density lipoprotein receptor |
| Tay-Sachs disease | Motor weakness, mental deterioration, death by age 3 yr | Hexosaminidase-A |
| Sickle-cell anemia | Pain, swelling in hands and feet; can lead to sudden severe pain in bones or joints and death | Hemoglobin |

The amino acid sequence of a protein is inextricably linked to its function. Proteins often contain crucial substructures within their amino acid sequence that are essential to their biological functions. The amino acid sequence in other regions might vary considerably without affecting these functions. The fraction of the sequence that is critical varies from protein to protein, complicating the task of relating sequence to structure, and structure to function. Before we can consider this problem further, however, we must examine how sequence information is obtained.

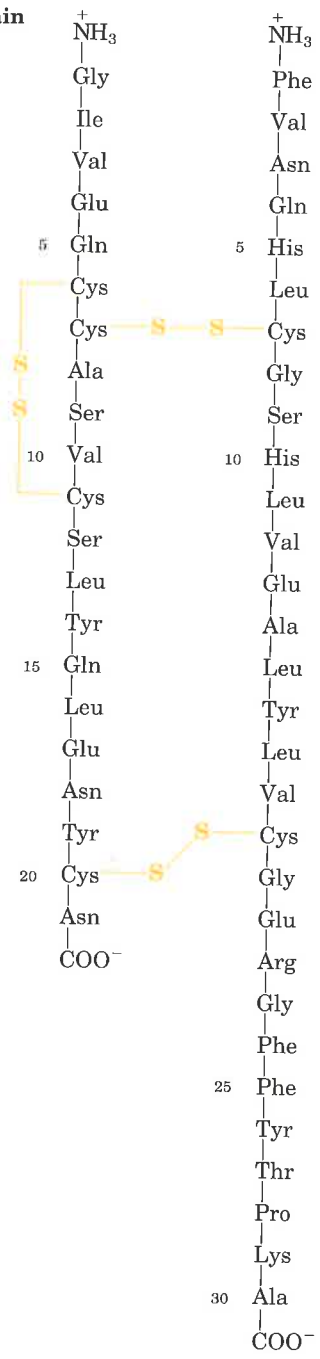



**Figure 6–10** The amino acid sequence of the two chains of bovine insulin, which are joined by disulfide cross-linkages. The A chain is identical in human, pig, dog, rabbit, and sperm whale insulins. The B chains of the cow, pig, dog, goat, and horse are identical. Such identities between similar proteins of different species are discussed later in this chapter.

## The Amino Acid Sequence of Polypeptide Chains Can Be Determined

Two major discoveries in 1953 ushered in the modern era of biochemistry. In that year James D. Watson and Francis Crick deduced the double-helical structure of DNA and proposed a structural basis for the precise replication of DNA (Chapter 12). Implicit in their proposal was the idea that the sequence of nucleotide units in DNA bears encoded genetic information. In that same year, Frederick Sanger worked out the sequence of amino acids in the polypeptide chains of the hormone insulin (Fig. 6–10), surprising many researchers who had long thought that elucidation of the amino acid sequence of a polypeptide would be a hopelessly difficult task. These achievements together suggested that the nucleotide sequence of DNA and the amino acid sequence of proteins were somehow related. Within just over a decade, the nucleotide code that determines the amino acid sequence of protein molecules had been revealed (Chapter 26).

Today the amino acid sequences of thousands of different proteins from many species are known, determined using principles first developed by Sanger. These methods are still in use, although with many variations and improvements in detail.

## Short Polypeptides Are Sequenced Using Automated Procedures

Three procedures are used in the determination of the sequence of a polypeptide chain (Fig. 6–11). The first is to hydrolyze it and determine its amino acid composition (Fig. 6–11a). This information is often valuable in later steps, and can also be useful in itself. Because amino acid composition differs from one protein to the next, it can serve as a kind of fingerprint. It can be used, for example, to help determine whether proteins isolated by different laboratories are the same or different.

Often, the next step is to identify the amino-terminal amino acid residue (Fig. 6–11b). For this purpose Sanger developed the reagent 1-fluoro-2,4-dinitrobenzene (FDNB; see Fig. 5–14). Other reagents used to label the amino-terminal residue are dansyl chloride and dabsyl chloride (see Figs. 5–14 and 5–18). The dansyl derivative is highly fluorescent and can be detected and measured in much lower concentrations than dinitrophenyl derivatives. The dabsyl derivative is intensely colored and also provides greater sensitivity than the dinitrophenyl compounds. These methods destroy the polypeptide and their utility is therefore limited to identification of the amino-terminal residue.

Frederick Sanger

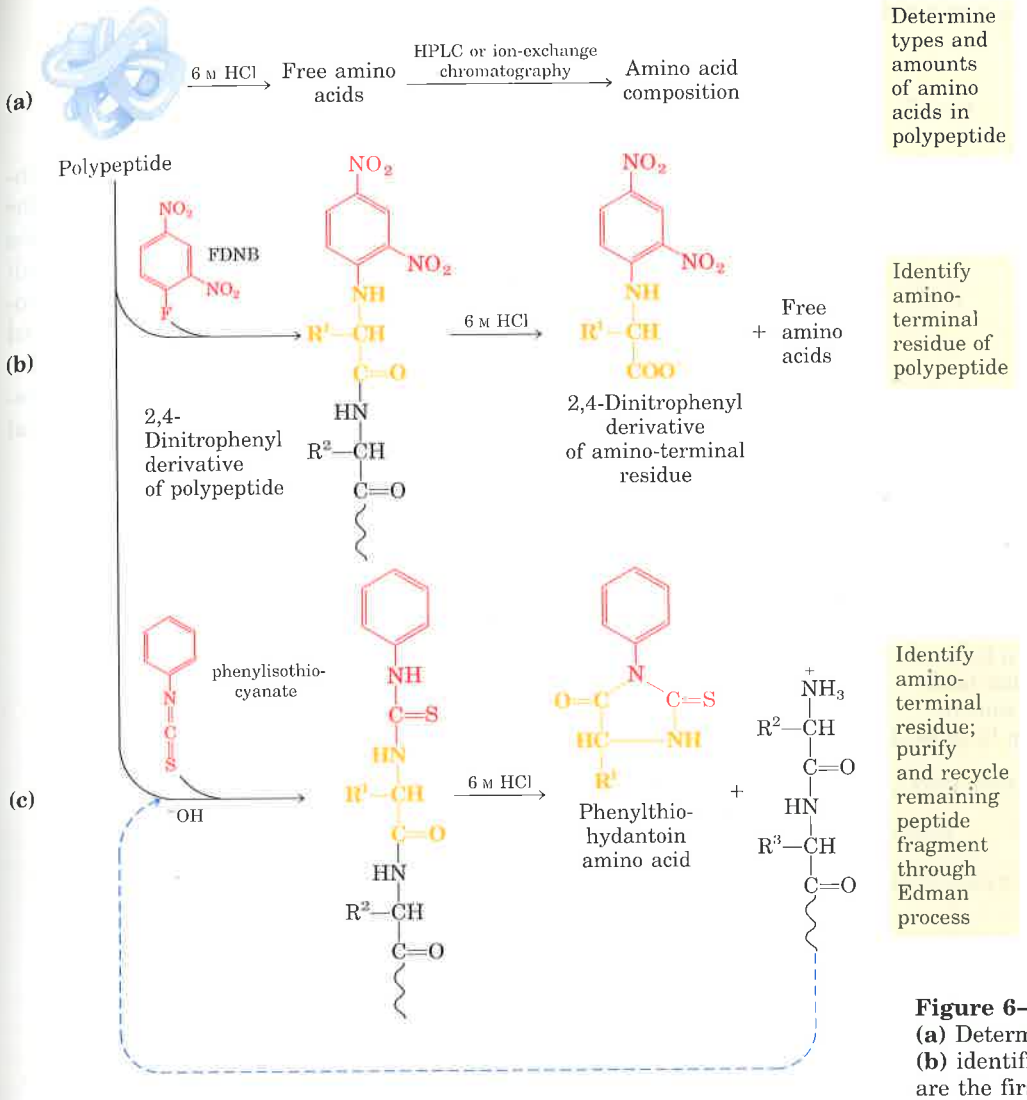



**Figure 6–11** Steps in sequencing a polypeptide. **(a)** Determination of amino acid composition and **(b)** identification of the amino-terminal residue are the first steps for many polypeptides. Sanger's method for identifying the amino-terminal residue is shown here. The Edman degradation procedure **(c)** reveals the entire sequence of a peptide. For shorter peptides, this method alone readily yields the entire sequence, and steps **(a)** and **(b)** are often omitted. The latter procedures are useful in the case of larger polypeptides, which are often fragmented into smaller peptides for sequencing (see Fig. 6–13).

To sequence the entire polypeptide, a chemical method devised by Pehr Edman is usually employed. The **Edman degradation** procedure labels and removes only the amino-terminal residue from a peptide, leaving all other peptide bonds intact (Fig. 6–11c). The peptide is reacted with phenylisothiocyanate, and the amino-terminal residue is ultimately removed as a phenylthiohydantoin derivative. After removal and identification of the amino-terminal residue, the *new* amino-terminal residue so exposed can be labeled, removed, and identified by repeating the same series of reactions. This procedure is repeated until the entire sequence is determined. Refinements of each step permit the sequencing of up to 50 amino acid residues in a large peptide.

The many individual steps and the careful bookkeeping required in the determination of the amino acid sequence of long polypeptide chains are usually carried out by programmed and automated analyzers. The Edman degradation is carried out on a programmed machine, called a **sequenator,** which mixes reagents in the proper proportions, separates the products, identifies them, and records the results. Such instruments have greatly reduced the time and labor required to determine the amino acid sequence of polypeptides. These methods are extremely sensitive. Often, less than a microgram of protein is sufficient to determine its complete amino acid sequence.

## Large Proteins Must Be Sequenced in Smaller Segments

The overall accuracy for determination of an amino acid sequence generally declines as the length of the polypeptide increases, especially for polypeptides longer than 50 amino acids. The very large polypeptides found in proteins must usually be broken down into pieces small enough to be sequenced efficiently. There are several steps in this process. First, any disulfide bonds are broken, and the protein is cleaved into a set of specific fragments by chemical or enzymatic methods. Each fragment is then purified, and sequenced by the Edman procedure. Finally, the order in which the fragments appear in the original protein is determined and disulfide bonds (if any) are located.

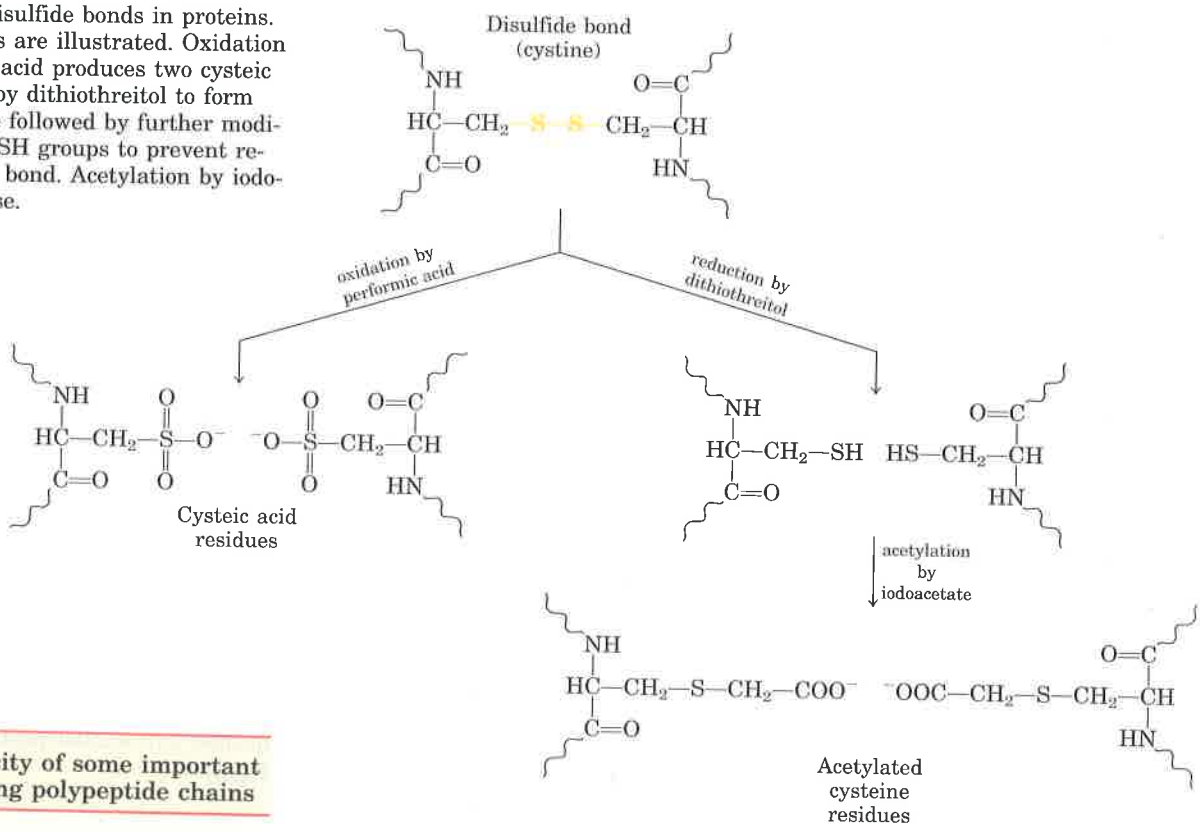



**Figure 6–12** Breaking disulfide bonds in proteins. The two common methods are illustrated. Oxidation of cystine with performic acid produces two cysteic acid residues. Reduction by dithiothreitol to form cysteine residues must be followed by further modification of the reactive —SH groups to prevent reformation of the disulfide bond. Acetylation by iodoacetate serves this purpose.

***Breaking Disulfide Bonds*** Disulfide bonds interfere with the sequencing procedure. A cystine residue (p. 116) that has one of its peptide bonds cleaved by the Edman procedure will remain attached to the polypeptide. Disulfide bonds also interfere with the enzymatic or chemical cleavage of the polypeptide (described below). Two approaches to irreversible breakage of disulfide bonds are outlined in Figure 6–12.

***Cleaving the Polypeptide Chain*** Several methods can be used for fragmenting the polypeptide chain. These involve a set of enzymes (proteases) and chemical reagents that cleave peptide chains adjacent to specific amino acid residues (Table 6–7). The digestive enzyme trypsin, for example, catalyzes the hydrolysis of only those peptide bonds in

**Table 6–7 The specificity of some important methods for fragmenting polypeptide chains**

| Treatment* | Cleavage points† |
|---|---|
| Trypsin | Lys, Arg (C) |
| *Submaxillarus* protease | Arg (C) |
| Chymotrypsin | Phe, Trp, Tyr (C) |
| *Staphylococcus aureus* V8 protease | Asp, Glu (C) |
| Asp-*N*-protease | Asp, Glu (N) |
| Pepsin | Phe, Trp, Tyr (N) |
| Cyanogen bromide | Met (C) |

* All of the enzymes or reagents listed are available from commercial sources.

† Residues furnishing the primary recognition point for the protease; peptide bond cleavage occurs either on the carbonyl (C) or amino (N) side of the indicated group of amino acids.

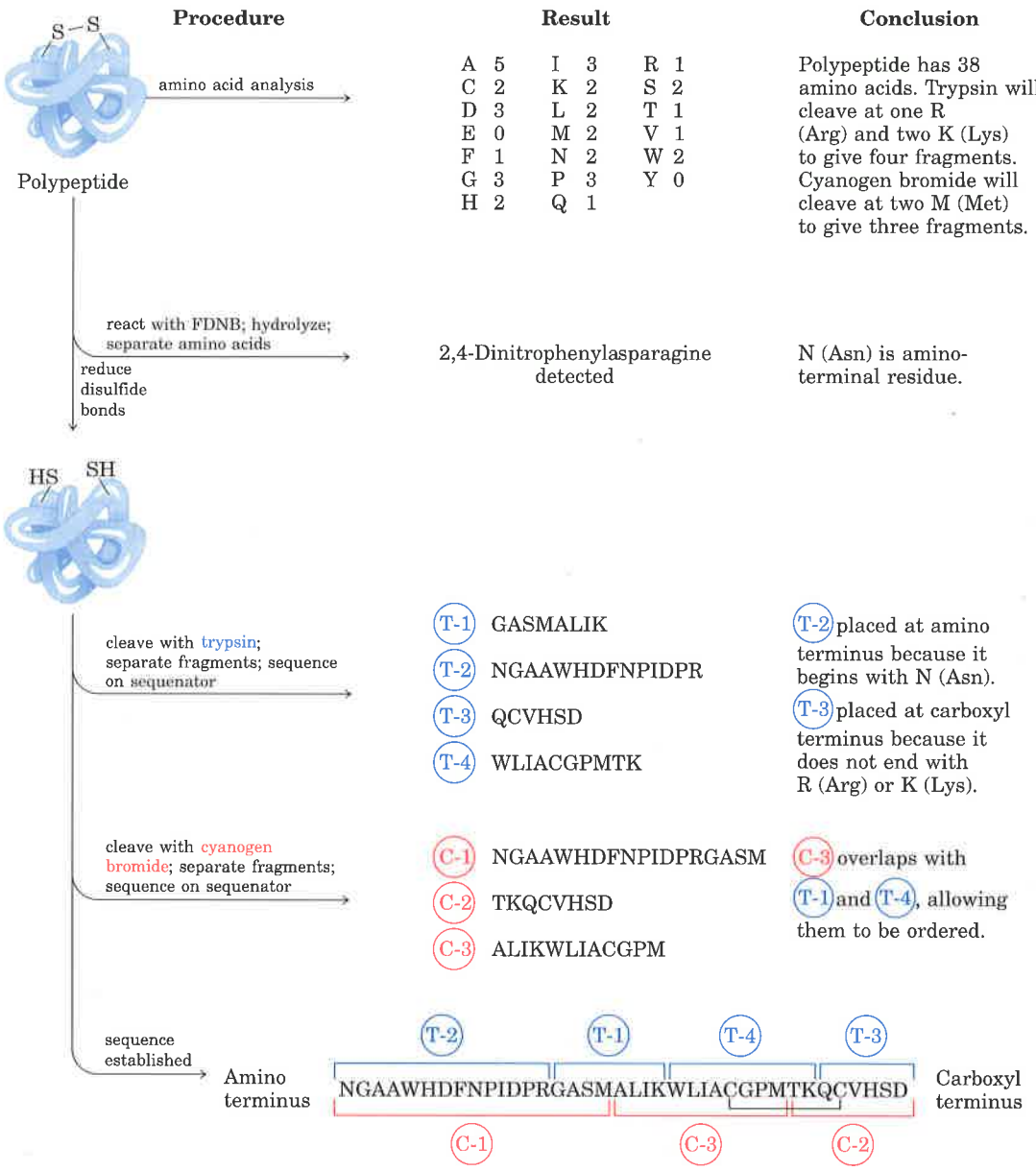



**Figure 6–13** Fragmenting proteins prior to sequencing, and placing peptide fragments in their proper order with overlaps. The one-letter abbreviations for amino acids are given in Table 5–1. In this example, there are only two Cys residues, thus one possibility for location of the disulfide bridge (black bracket). In polypeptides with three or more Cys residues, disulfide bridges can be located as described in the text.

which the carbonyl group is contributed by either a Lys or an Arg residue, regardless of the length or amino acid sequence of the chain. The number of smaller peptides produced by trypsin cleavage can thus be predicted from the total number of Lys or Arg residues in the original polypeptide (Fig. 6–13). A polypeptide with five Lys and/or Arg residues will usually yield six smaller peptides on cleavage with trypsin. Moreover, all except one of these will have a carboxyl-terminal Lys or Arg. The fragments produced by trypsin action are separated by chromatographic or electrophoretic methods.

***Sequencing of Peptides*** All the peptide fragments resulting from the action of trypsin are sequenced separately by the Edman procedure.

***Ordering Peptide Fragments*** The order of these trypsin fragments in the original polypeptide chain must now be determined. Another sample of the intact polypeptide is cleaved into small fragments using a different enzyme or reagent, one that cleaves peptide bonds at points other than those cleaved by trypsin. For example, the reagent cyanogen bromide cleaves only those peptide bonds in which the carbonyl group is contributed by Met (Table 6–7). The fragments resulting from this new procedure are then separated and sequenced as before.

The amino acid sequences of each fragment obtained by the two cleavage procedures are examined, with the objective of finding peptides from the second procedure whose sequences establish continuity, because of overlaps, between the fragments obtained by the first cleavage procedure (Fig. 6–13). Overlapping peptides obtained from the second fragmentation yield the correct order of the peptide fragments produced in the first. Moreover, the two sets of fragments can be compared for possible errors in determining the amino acid sequence of each fragment. If the amino-terminal amino acid has been identified before the original cleavage of the protein, this information can be used to establish which fragment is derived from the amino terminus.

If the second cleavage procedure fails to establish continuity between all peptides from the first cleavage, a third or even a fourth cleavage method must be used to obtain a set of peptides that can provide the necessary overlap(s). A variety of proteolytic enzymes with different specificities are available (Table 6–7).

***Locating Disulfide Bonds*** After sequencing is completed, locating the disulfide bonds requires an additional step. A sample of the protein is again cleaved with a reagent such as trypsin, this time without first breaking the disulfide bonds. When the resulting peptides are separated by electrophoresis and compared with the original set of peptides generated by trypsin, two of the original peptides will be missing and a new, larger peptide will appear. The two missing peptides represent the regions of the intact polypeptide that are linked by a disulfide bond.

## Amino Acid Sequences Can Be Deduced from DNA Sequences

The approach outlined above is not the only way to obtain amino acid sequences. The development of rapid DNA sequencing methods (Chapter 12), the elucidation of the genetic code (Chapter 26), and the development of techniques for the isolation of genes (Chapter 28) make it possible to deduce the sequence of a polypeptide by determining the sequence of nucleotides in its gene (Fig. 6–14). The two techniques are complementary. When the gene is available, sequencing the DNA can be faster and more accurate than sequencing the protein. If the gene has not been isolated, direct sequencing of peptides is necessary, and this can provide information (e.g., the location of disulfide bonds) not available in a DNA sequence. In addition, a knowledge of the amino acid sequence can greatly facilitate the isolation of the corresponding gene (Chapter 28).

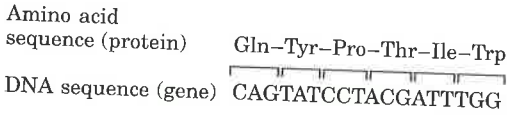



**Figure 6–14** Correspondence of DNA and amino acid sequences. Each amino acid is encoded by a specific sequence of three nucleotides (triplet) in DNA. The genetic code is described in detail in Chapter 26.

## Amino Acid Sequences Provide Important Biochemical Information

The sequence of amino acids in a protein can offer insights into its three-dimensional structure and its function, cellular location, and evolution. Most of these insights are derived by searching for similarities with other known sequences. Thousands of sequences are known and available in computerized data bases. The comparison of a newly obtained sequence with this large bank of stored sequences often reveals relationships both surprising and enlightening.

The relationship between amino acid sequence and three-dimensional structure, and between structure and function, is not understood in detail. However, a growing number of protein families are being revealed that have at least some shared structural and functional features that can be readily identified on the basis of amino acid sequence similarities alone. For example, there are four major families of proteases, several families of naturally occurring protease inhibitors, a large number of closely related protein kinases, and a similar large number of related protein phosphatases. Individual proteins are generally assigned to families by the degree of similarity in amino acid sequence (identical to other members of the family across 30% or more of the sequence), and proteins in these families generally share at least some structural and functional characteristics. Some families are defined, however, by identities involving only a few amino acids that are critical to a certain function. Many membrane-bound protein receptors share important structural features and have similar amino acid sequences, even though the extracellular molecules they bind are quite different. Even the immunoglobulin family includes a host of extracellular and cell-surface proteins in addition to antibodies.

The similarities may involve the entire protein or may be confined to relatively small segments of it. A number of similar substructures (domains) occur in many functionally unrelated proteins. An example is a 40 to 45 amino acid sequence called the EGF (epidermal growth factor) domain that makes up part of the structure of urokinase, the low-density lipoprotein receptor, several proteins involved in blood clotting, and many others. These domains often fold up into structural configurations that have an unusual degree of stability or that are specialized for a certain environment. Evolutionary relationships can also be inferred from the structural and functional similarities within protein families.

Certain amino acid sequences often serve as signals that determine the cellular location, chemical modification, and half-life of a protein. Special signal sequences, usually at the amino terminus, are used to target certain proteins for export from the cell, while other proteins are distributed to the nucleus, the cell surface, the cytosol, and other cellular locations. Other sequences act as attachment sites for prosthetic groups, such as glycosyl groups in glycoproteins and lipids in lipoproteins. Some of these signals are well characterized, and are easily recognized if they occur in the sequence of a newly discovered protein.

The probability that information about a new protein can be deduced from its primary structure improves constantly with the almost daily addition to the number of published amino acid sequences stored in shared databanks.

## Homologous Proteins from Different Species Have Homologous Sequences

Several important conclusions have come from study of the amino acid sequences of homologous proteins from different species. **Homologous proteins** are those that are evolutionarily related. They usually perform the same function in different species; an example is hemoglobin, which has the same oxygen-transport function in different vertebrates. Homologous proteins from different species often have polypeptide chains that are identical or nearly identical in length. Many positions in the amino acid sequence are occupied by the same amino acid in all species and are thus called **invariant residues.** But in other positions there may be considerable variation in the amino acid from one species to another; these are called **variable residues.**

The functional significance of sequence homology can be illustrated by **cytochrome *c*,** an iron-containing mitochondrial protein that transfers electrons during biological oxidations in eukaryotic cells. The polypeptide chain of this protein has a molecular weight of about 13,000 and has about 100 amino acid residues in most species. The amino acid sequences of cytochrome *c* from over 60 different species have been determined, and 27 positions in the chain of amino acid residues are invariant in all species tested (Fig. 6–15), suggesting that they are the most important residues specifying the biological activity of cytochrome *c*. The residues in other positions in the chain exhibit some interspecies variation. There are clear gradations in the number of changes observed in the variable residues. In some positions, all substitutions involve similar amino acid residues (e.g., Arg will replace Lys, both of which are positively charged); these are called **conservative substitutions.** At other positions the substitutions are more random. As we will show in the next chapter, the polypeptide chains of proteins are folded into characteristic and specific conformations and these conformations depend on amino acid sequence. Clearly, the invariant residues are more critical to the structure and function of a protein than the variable ones. Recognizing which amino acids fall into each category is an important step in deciphering the complicated question of how amino acid sequence is translated into a specific three-dimensional structure.

The variable amino acids provide information of another sort. Evolution is sometimes regarded as a theory that is accepted but difficult to test, yet the phylogenetic trees established by taxonomy have been tested and experimentally confirmed through biochemistry. The exam-

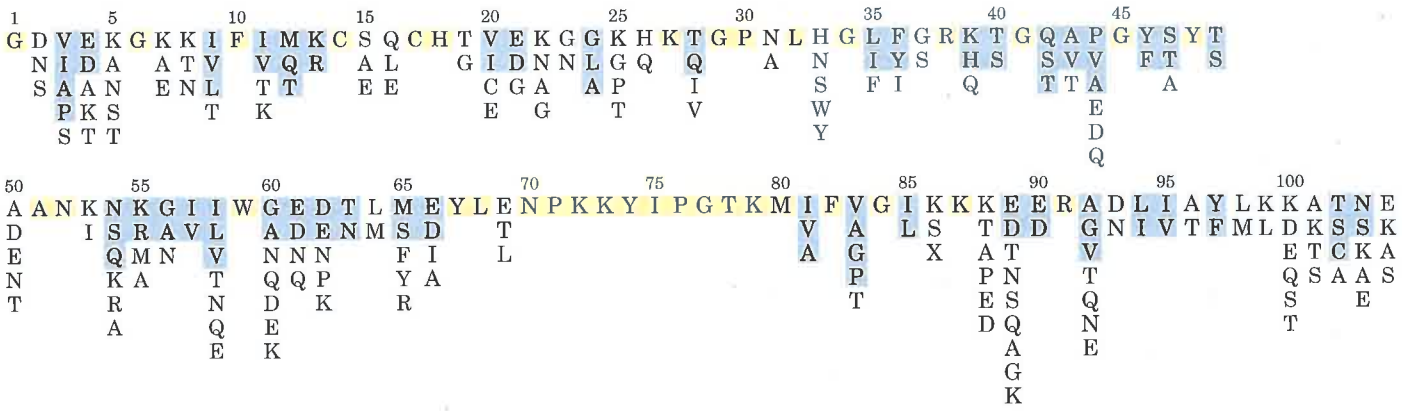



**Figure 6–15** The amino acid sequence of human cytochrome *c*. Amino acid substitutions found at different positions in the cytochrome *c* of other species are listed below the sequence of the human protein. The amino acids are color-coded to help distinguish conservative and nonconservative substitutions: invariant amino acids are shaded in yellow, conservative amino acid substitutions are shaded in blue, and nonconservative substitutions are unshaded. X is an unusual amino acid, trimethyllysine. The one-letter abbreviations for amino acids are used here (see Table 5–1).

ination of sequences of cytochrome *c* and other homologous proteins has led to an important conclusion: the number of residues that differ in homologous proteins from any two species is in proportion to the phylogenetic difference between those species. For example, 48 amino acid residues differ in the cytochrome *c* molecules of the horse and of yeast, which are very widely separated species, whereas only two residues differ in the cytochrome *c* of the much more closely related duck and chicken. In fact, the cytochrome *c* molecule has identical amino acid sequences in the chicken and the turkey, and in the pig, cow, and sheep. Information on the number of residue differences between homologous proteins of different species allows the construction of evolutionary maps that show the origin and sequence of development of different animals and plants during the evolution of species (Fig. 6–16). The relationships established by taxonomy and biochemistry agree well.

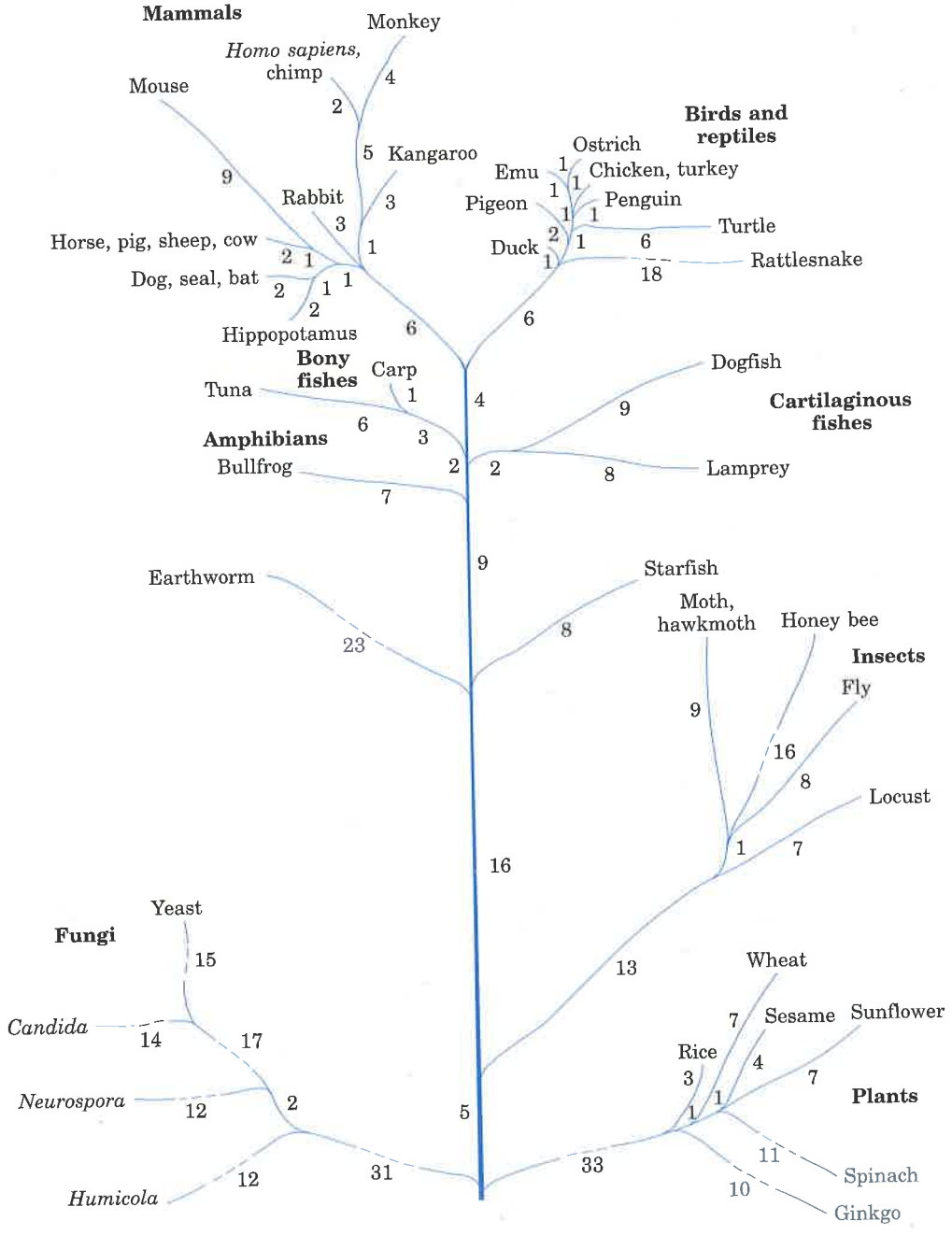



**Figure 6–16** Main branches of the evolutionary tree constructed from the number of amino acid differences between cytochrome *c* molecules of different species. The numbers represent the number of residues by which the cytochrome *c* of a given line of organism differs from its ancestors.

## Summary

Cells generally contain thousands of different proteins, each with a different function or biological activity. These functions include enzymatic catalysis, molecular transport, nutrition, cell or organismal motility, structural roles, organismal defense, regulation, and many others. Proteins consist of very long polypeptide chains having from 100 to over 2,000 amino acid residues joined by peptide linkages. Some proteins have several polypeptide chains, which are then referred to as subunits. Simple proteins yield only amino acids on hydrolysis; conjugated proteins contain in addition some other component, such as a metal ion or organic prosthetic group.

Proteins are purified by taking advantage of properties in which they differ, such as size, shape, binding affinities, charge, etc. Purification also requires a method for quantifying or assaying a particular protein in the presence of others. Proteins can be both separated and visualized by electrophoretic methods. Antibodies that specifically bind a certain protein can be used to detect and locate that protein in a solution, a gel, or even in the interior of a cell.

All proteins are made from the same set of 20 amino acids. Their differences in function result from differences in the composition and sequence of their amino acids. The amino acid sequences of polypeptide chains can be established by fragmenting them into smaller pieces using several specific reagents, and determining the amino acid sequence of each fragment by the Edman degradation procedure. The sequencing of suitably sized peptide fragments has been automated. The peptide fragments are then placed in the correct order by finding sequence overlaps between fragments generated by different methods. Protein sequences can also be deduced from the nucleotide sequence of the corresponding gene in the DNA. The amino acid sequence can be compared with the thousands of known sequences, often revealing insights into the structure, function, cellular location, and evolution of the protein.

Homologous proteins from different species show sequence homology: certain positions in the polypeptide chains contain the same amino acids, regardless of the species. In other positions the amino acids may differ. The invariant residues are evidently essential to the function of the protein. The degree of similarity between amino acid sequences of homologous proteins from different species correlates with the evolutionary relationship of the species.

## Further Reading

See Chapter 5 for additional useful references.

### *Properties of Proteins*

**Creighton, T.E.** (1984) *Proteins: Structures and Molecular Properties,* W.H. Freeman and Company, New York.

**Dickerson, R.E. & Geis, I.** (1983) *Proteins: Structure, Function, and Evolution,* 2nd edn, The Benjamin/Cummings Publishing Company, Menlo Park, CA.

*A beautifully illustrated introduction to proteins.*

**Doolittle, R.F.** (1985) Proteins. *Sci. Am.* **253** (October), 88–99.

*An overview that highlights evolutionary relationships.*

**Srinavasan, P.R., Fruton, J.S., & Edsall, J.T. (eds)** (1979) The Origins of Modern Biochemistry: A Retrospect on Proteins. *Ann. N.Y. Acad. Sci.* **325.**

*A collection of very interesting articles on the history of protein research.*

Structure and Function of Proteins. (1989) *Trends Biochem. Sci.* **14** (July).

*A special issue devoted to reviews on protein chemistry and protein structure.*

### *Working with Proteins*

**Hirs, C.H.W. & Timasheff, S.N. (eds)** (1983) *Methods in Enzymology,* Vol. 91, Part I: *Enzyme Structure,* Academic Press, Inc., New York.

*An excellent collection of authoritative articles on techniques in protein chemistry. Includes information on sequencing.*

**Kornberg, A.** (1990) Why purify enzymes? In *Methods in Enzymology,* Vol. 182: *Guide to Protein Purification* (Deutscher, M.P., ed), pp. 1–5, Academic Press, Inc., New York.

*The critical role of classical biochemical methods in a new age.*

**O'Farrell, P.H.** (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250,** 4007–4021.

*An interesting attempt to count all the proteins in the* E. coli *cell.*

**Plummer, David T.** (1987) *An Introduction to Practical Biochemistry,* 3rd edn, McGraw-Hill, London.

*Good descriptions of many techniques for beginning students.*

**Scopes, R.K.** (1987) *Protein Purification: Principles and Practice,* 2nd edn, Springer-Verlag, New York.

**Tonegawa, S.** (1985) The molecules of the immune system. *Sci. Am.* **253** (October), 122–131.

### *The Covalent Structure of Proteins*

**Dickerson, R.E.** (1972) The structure and history of an ancient protein. *Sci. Am.* **226** (April), 58–72.

*A nice summary of information gleaned from interspecies comparisons of cytochrome* c *sequences.*

**Doolittle, R.** (1981) Similar amino acid sequences: chance or common ancestry? *Science* **214,** 149–159.

*A good discussion of what can be learned by comparing amino acid sequences.*

**Hunkapiller, M.W., Strickler, J.E., & Wilson, K.J.** (1984) Contemporary methodology for protein structure determination. *Science* **226,** 304–311.

**Reidhaar-Olson, J.F. & Sauer, R.T.** (1988) Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences. *Science* **241,** 53–57.

*A systematic study of possible amino acid substitutions in a short segment of one protein.*

**Wilson, A.C.** (1985) The molecular basis of evolution. *Sci. Am.* **253** (October), 164–173.

## Problems

**1.** *How Many β-Galactosidase Molecules Are Present in an* E. coli *Cell? E. coli* is a rod-shaped bacterium 2 μm long and 1 μm in diameter. When grown on lactose (a sugar found in milk), the bacterium synthesizes the enzyme β-galactosidase ($M_r$ 450,000), which catalyzes the breakdown of lactose. The average density of the bacterial cell is 1.2 g/mL, and 14% of its total mass is soluble protein, of which 1.0% is β-galactosidase. Calculate the number of β-galactosidase molecules in an *E. coli* cell grown on lactose.

**2.** *The Number of Tryptophan Residues in Bovine Serum Albumin* A quantitative amino acid analysis reveals that bovine serum albumin contains 0.58% by weight of tryptophan, which has a molecular weight of 204.

(a) Calculate the minimum molecular weight of bovine serum albumin (i.e., assuming there is only one tryptophan residue per protein molecule).

(b) Gel filtration of bovine serum albumin gives a molecular weight estimate of about 70,000. How many tryptophan residues are present in a molecule of serum albumin?

**3.** *The Molecular Weight of Ribonuclease* Lysine makes up 10.5% of the weight of ribonuclease. Calculate the minimum molecular weight of ribonuclease. The ribonuclease molecule contains ten lysine residues. Calculate the molecular weight of ribonuclease.

**4.** *The Size of Proteins* What is the approximate molecular weight of a protein containing 682 amino acids in a single polypeptide chain?

**5.** *Net Electric Charge of Peptides* A peptide isolated from the brain has the sequence

Glu–His–Trp–Ser–Tyr–Gly–Leu–Arg–Pro–Gly

Determine the net charge on the molecule at pH 3. What is the net charge at pH 5.5? At pH 8? At pH 11? Estimate the pI for this peptide. (Use $pK_a$ values for side chains and terminal amino and carboxyl groups as given in Table 5–1.)

**6.** *The Isoelectric Point of Pepsin* Pepsin of gastric juice (pH ≈ 1.5) has a pI of about 1, much lower than that of other proteins (see Table 6–5). What functional groups must be present in relatively large numbers to give pepsin such a low pI? What amino acids can contribute such groups?

**7.** *The Isoelectric Point of Histones* Histones are proteins of eukaryotic cell nuclei. They are tightly

bound to deoxyribonucleic acid (DNA), which has many phosphate groups. The pI of histones is very high, about 10.8. What amino acids must be present in relatively large numbers in histones? In what way do these residues contribute to the strong binding of histones to DNA?

**8.** *Solubility of Polypeptides* One method for separating polypeptides makes use of their differential solubilities. The solubility of large polypeptides in water depends upon the relative polarity of their R groups, particularly on the number of ionized groups: the more ionized groups there are, the more soluble the polypeptide. Which of each pair of polypeptides below is more soluble at the indicated pH?

(a) $(Gly)_{20}$ or $(Glu)_{20}$ at pH 7.0

(b) $(Lys–Ala)_3$ or $(Phe–Met)_3$ at pH 7.0

(c) $(Ala–Ser–Gly)_5$ or $(Asn–Ser–His)_5$ at pH 6.0

(d) $(Ala–Asp–Gly)_5$ or $(Asn–Ser–His)_5$ at pH 3.0

**9.** *Purification of an Enzyme* A biochemist discovers and purifies a new enzyme, generating the purification table below:

| Procedure | Total protein (mg) | Activity (units) |
|---|---|---|
| 1. Crude extract | 20,000 | 4,000,000 |
| 2. Precipitation (salt) | 5,000 | 3,000,000 |
| 3. Precipitation (pH) | 4,000 | 1,000,000 |
| 4. Ion-exchange chromatography | 200 | 800,000 |
| 5. Affinity chromatography | 50 | 750,000 |
| 6. Size-exclusion chromatography | 45 | 675,000 |

(a) From the information given in the table, calculate the specific activity of the enzyme solution after each purification procedure.

(b) Which of the purification procedures used for this enzyme is most effective (i.e., gives the greatest increase in purity)?

(c) Which of the purification procedures is least effective?

(d) Is there any indication in this table that the enzyme is now pure? What else could be done to estimate the purity of the enzyme preparation?

**10.** *Fragmentation of a Polypeptide Chain by Proteolytic Enzymes* Trypsin and chymotrypsin are specific enzymes that catalyze the hydrolysis of polypeptides at specific locations (Table 6–7). The sequence of the B chain of insulin is shown below. Note that the cystine cross-linkage between the A and B chains has been cleaved through the action of performic acid (see Fig. 6–12).

Phe–Val–Asn–Gln–His–Leu–$CysSO_3^-$–Gly–Ser–His–Leu–Val–Glu–Ala–Leu–Tyr–Leu–Val–$CysSO_3^-$–Gly–Glu–Arg–Gly–Phe–Phe–Tyr–Thr–Pro–Lys–Ala

Indicate the points in the B chain that are cleaved by (a) trypsin and (b) chymotrypsin. Note that these proteases will not remove single amino acids from either end of a polypeptide chain.

**11.** *Sequence Determination of the Brain Peptide Leucine Enkephalin* A group of peptides that influence nerve transmission in certain parts of the brain has been isolated from normal brain tissue. These peptides are known as opioids, because they bind to specific receptors that bind opiate drugs, such as morphine and naloxone. Opioids thus mimic some of the properties of opiates. Some researchers consider these peptides to be the brain's own pain killers. Using the information below, determine the amino acid sequence of the opioid leucine enkephalin. Explain how your structure is consistent with each piece of information.

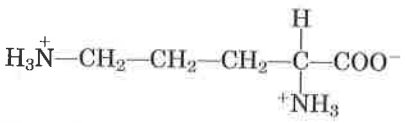
(a) Complete hydrolysis by 1 M HCl at 110 °C followed by amino acid analysis indicated the presence of Gly, Leu, Phe, and Tyr, in a 2:1:1:1 molar ratio.

(b) Treatment of the peptide with 1-fluoro-2,4-dinitrobenzene followed by complete hydrolysis and chromatography indicated the presence of the 2,4-dinitrophenyl derivative of tyrosine. No free tyrosine could be found.

(c) Complete digestion of the peptide with pepsin followed by chromatography yielded a dipeptide containing Phe and Leu, plus a tripeptide containing Tyr and Gly in a 1:2 ratio.

**12.** *Structure of a Peptide Antibiotic from* Bacillus brevis Extracts from the bacterium *Bacillus brevis* contain a peptide with antibiotic properties. Such peptide antibiotics form complexes with metal ions and apparently disrupt ion transport across the cell membrane, killing certain bacterial species. The structure of the peptide has been determined from the following observations.

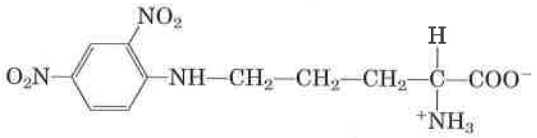
(a) Complete acid hydrolysis of the peptide followed by amino acid analysis yielded equimolar amounts of Leu, Orn, Phe, Pro, and Val. Orn is ornithine, an amino acid not present in proteins but present in some peptides. It has the structure



(b) The molecular weight of the peptide was estimated as about 1,200.

(c) When treated with the enzyme carboxypeptidase, the peptide failed to undergo hydrolysis.

(d) Treatment of the intact peptide with 1-fluoro-2,4-dinitrobenzene, followed by complete hydrolysis and chromatography, yielded only free amino acids and the following derivative:



(Hint: Note that the 2,4-dinitrophenyl derivative involves the amino group of a side chain rather than the $\alpha$-amino group.)

(e) Partial hydrolysis of the peptide followed by chromatographic separation and sequence analysis yielded the di- and tripeptides below (the amino-terminal amino acid is always at the left):

Leu–Phe Phe–Pro Orn–Leu Val–Orn
Val–Orn–Leu Phe–Pro–Val Pro–Val–Orn

Given the above information, deduce the amino acid sequence of the peptide antibiotic. Show your reasoning. When you have arrived at a structure, go back and demonstrate that it is consistent with *each* experimental observation.

CHAPTER

# 7

# The Three-Dimensional Structure of Proteins

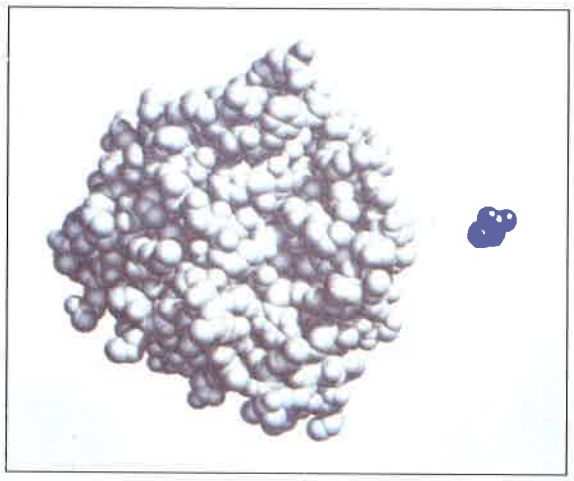


**Figure 7–1** The structure of the enzyme chymotrypsin, a globular protein. A molecule of glycine (blue) is shown for size comparison.

The covalent backbone of proteins is made up of hundreds of individual bonds. If free rotation were possible around even a fraction of these bonds, proteins could assume an almost infinite number of three-dimensional structures. Each protein has a specific chemical or structural function, however, strongly suggesting that each protein has a unique three-dimensional structure (Fig. 7–1). The simple fact that proteins can be crystallized provides strong evidence that this is the case. The ordered arrays of molecules in a crystal can generally form only if the molecular units making up the crystal are identical. The enzyme urease ($M_r$ 483,000) was among the first proteins crystallized, by James Sumner in 1926. This accomplishment demonstrated dramatically that even very large proteins are discrete chemical entities with unique structures, and it revolutionized thinking about proteins.

In this chapter, we will explore the three-dimensional structure of proteins, emphasizing several principles. First, the three-dimensional structure of a protein is determined by its amino acid sequence. Second, the function of a protein depends upon its three-dimensional structure. Third, the three-dimensional structure of a protein is unique, or nearly so. Fourth, the most important forces stabilizing the specific three-dimensional structure maintained by a given protein are noncovalent interactions. Finally, even though the structure of proteins is complicated, several common patterns can be recognized.

The relationship between the amino acid sequence and the three-dimensional structure of a protein is an intricate puzzle that has yet to be solved in detail. Polypeptides with very different amino acid sequences sometimes assume similar structures, and similar amino acid sequences sometimes yield very different structures. To find and understand patterns in this biochemical labyrinth requires a renewed appreciation for fundamental principles of chemistry and physics.

## Overview of Protein Structure

The spatial arrangement of atoms in a protein is called a **conformation.** The term conformation refers to a structural state that can, without breaking any covalent bonds, interconvert with other structural states. A change in conformation could occur, for example, by rotation about single bonds. Of the innumerable conformations that are theoretically possible in a protein containing hundreds of single bonds, one generally predominates. This is usually the conformation that is ther-

modynamically the most stable, having the lowest Gibbs' free energy ($G$). Proteins in their functional conformation are called **native** proteins.

What principles determine the most stable conformation of a protein? Although protein structures can seem hopelessly complex, close inspection reveals recurring structural patterns. The patterns involve different levels of structural complexity, and we now turn to a biochemical convention that serves as a framework for much of what follows in this chapter.

## There Are Four Levels of Architecture in Proteins

Conceptually, protein structure can be considered at four levels (Fig. 7–2). **Primary structure** includes all the covalent bonds between amino acids and is normally defined by the sequence of peptide-bonded amino acids and locations of disulfide bonds. The relative spatial arrangement of the linked amino acids is unspecified.

Polypeptide chains are not free to take up any three-dimensional structure at random. Steric constraints and many weak interactions stipulate that some arrangements will be more stable than others. **Secondary structure** refers to regular, recurring arrangements in space of adjacent amino acid residues in a polypeptide chain. There are a few common types of secondary structure, the most prominent being the $\alpha$ helix and the $\beta$ conformation. **Tertiary structure** refers to the spatial relationship among all amino acids in a polypeptide; it is the complete three-dimensional structure of the polypeptide. The boundary between secondary and tertiary structure is not always clear. Several different types of secondary structure are often found within the three-dimensional structure of a large protein. Proteins with several polypeptide chains have one more level of structure: **quaternary structure,** which refers to the spatial relationship of the polypeptides, or subunits, within the protein.

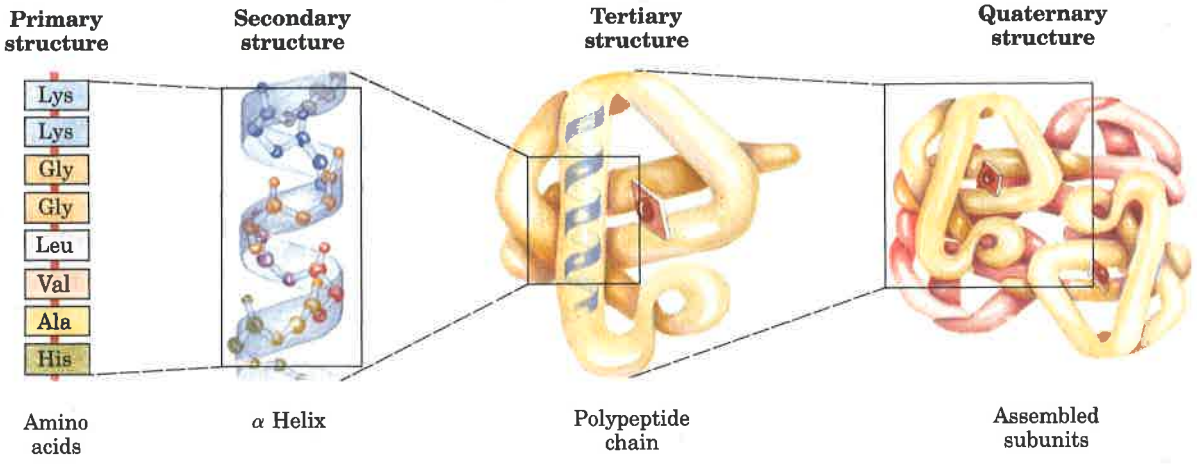



**Figure 7–2** Levels of structure in proteins. The *primary structure* consists of a sequence of amino acids linked together by covalent peptide bonds, and includes any disulfide bonds. The resulting polypeptide can be coiled into an $\alpha$ helix, one form of *secondary structure*. The helix is a part of the *tertiary structure* of the folded polypeptide, which is itself one of the subunits that make up the *quaternary structure* of the multimeric protein, in this case hemoglobin.

Continued advances in the understanding of protein structure, folding, and evolution have made it necessary to define two additional structural levels intermediate between secondary and tertiary structure. A stable clustering of several elements of secondary structure is sometimes referred to as **supersecondary structure.** The term is used to describe particularly stable arrangements that occur in many

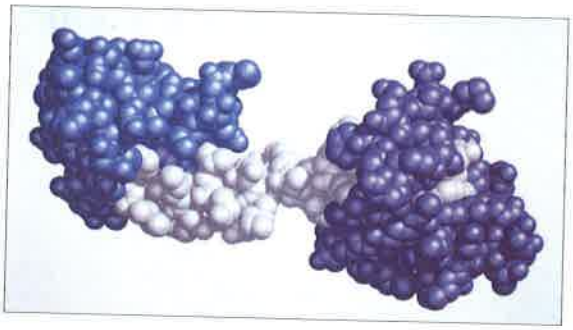


**Figure 7–3** The different structural domains in the polypeptide troponin C, a calcium-binding protein associated with muscle. The separate calcium-binding domains, indicated in blue and purple, are connected by a long $\alpha$ helix, shown in white.

different proteins and sometimes many times in a single protein. A somewhat higher level of structure is the **domain.** This refers to a compact region, including perhaps 40 to 400 amino acids, that is a distinct structural unit within a larger polypeptide chain. A polypeptide that is folded into a dumbbell-like shape might be considered to have two domains, one at either end. Many domains fold independently into thermodynamically stable structures. A large polypeptide chain can contain several domains that often are readily distinguishable within the overall structure (Fig. 7–3). In some cases the individual domains have separate functions. As we will see, important patterns exist at each of these levels of structure that provide clues to understanding the overall structure of large proteins.

## A Protein's Conformation Is Stabilized Largely by Weak Interactions

The native conformation of a protein is only marginally stable; the difference in free energy between the folded and unfolded states in typical proteins under physiological conditions is in the range of only 20 to 65 kJ/mol. A given polypeptide chain can theoretically assume countless different conformations, and as a result the unfolded state of a protein is characterized by a high degree of conformational entropy. This entropy, and the hydrogen-bonding interactions of many groups in the polypeptide chain with solvent (water), tend to maintain the unfolded state. The chemical interactions that counteract these effects and stabilize the native conformation include disulfide bonds and the weak (noncovalent) interactions described in Chapter 4: hydrogen bonds, and hydrophobic, ionic, and van der Waals interactions. An appreciation of the role of these weak interactions is especially important to understanding how polypeptide chains fold into specific secondary, tertiary, and quaternary structures.

Every time a bond is formed between two atoms, some free energy is released in the form of heat or entropy. In other words, the formation of bonds is accompanied by a favorable (negative) change in free energy. The $\Delta G$ for covalent bond formation is generally in the range of $-200$ to $-460$ kJ/mol. For weak interactions, $\Delta G = -4$ to $-30$ kJ/mol. Although covalent bonds are clearly much stronger, weak interactions predominate as a stabilizing force in protein structure because of their number. In general, the protein conformation with the lowest free energy (i.e., the most stable) is the one with the maximum number of weak interactions.

The stability of a protein is not simply the sum of the free energies of formation of the many weak interactions within it, however. We have already noted that the stability of proteins is marginal. Every hydrogen-bonding group in a polypeptide chain was hydrogen bonded to water prior to folding. For every hydrogen bond formed in a protein, hydrogen bonds (of similar strength) between the same groups and water were broken. The net stability contributed by a given weak interaction, or the *difference* in free energies of the folded and unfolded state, is close to zero. We must therefore explain why the native conformation of a protein is favored. The contribution of weak interactions to protein stability can be understood in terms of the properties of water (Chapter 4). Pure water contains a network of hydrogen-bonded water molecules. No other molecule has the hydrogen-bonding potential of water, and other molecules present in an aqueous solution will disrupt

the hydrogen bonding of water to some extent. Optimizing the hydrogen bonding of water around a hydrophobic molecule results in the formation of a highly structured shell or solvation layer of water in the immediate vicinity, resulting in an unfavorable decrease in the entropy of water. The association among hydrophobic or nonpolar groups results in a decrease in this structured solvation layer, or a favorable increase in entropy. As described in Chapter 4, this entropy term is the major thermodynamic driving force for the association of hydrophobic groups in aqueous solution, and hydrophobic amino acid side chains therefore tend to be clustered in a protein's interior, away from water.

The formation of hydrogen bonds and ionic interactions in a protein is also driven largely by this same entropic effect. Polar groups can generally form hydrogen bonds with water and hence are soluble in water. However, the number of hydrogen bonds per unit mass is generally greater for pure water than for any other liquid or solution, and there are limits to the solubility of even the most polar molecules because of the net decrease in hydrogen bonding that occurs when they are present. Therefore, a solvation shell of structured water will also form to some extent around polar molecules. Even though the energy of formation of an intramolecular hydrogen bond or ionic interaction between two polar groups in a macromolecule is largely canceled out by the elimination of such interactions between the same groups and water, the release of structured water when the intramolecular interaction is formed provides an entropic driving force for folding. Most of the net change in free energy that occurs when weak interactions are formed within a protein is therefore derived from the increase in entropy in the surrounding aqueous solution.

Of the different types of weak interactions, hydrophobic interactions are particularly important in stabilizing a protein conformation; the interior of a protein is generally a densely packed core of hydrophobic amino acid side chains. It is also important that any polar or charged groups in the protein interior have suitable partners for hydrogen bonding or ionic interactions. One hydrogen bond makes only a small apparent contribution to the stability of a native structure, but the presence of a single hydrogen-bonding group without a partner in the hydrophobic core of a protein can be so *destabilizing* that conformations containing such a group are often thermodynamically untenable.

Most of the structural patterns outlined in this chapter reflect these two simple rules: (1) hydrophobic residues must be buried in the protein interior and away from water, and (2) the number of hydrogen bonds must be maximized. Insoluble proteins and proteins within membranes (Chapter 10) follow somewhat different rules because of their function or their environment, but weak interactions are still critical structural elements.

Linus Pauling

Robert Corey
1897–1971

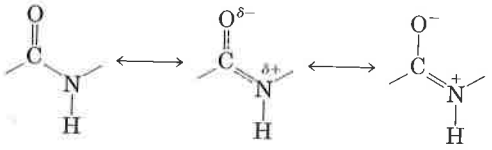
## Protein Secondary Structure

Several types of secondary structure are particularly stable and occur widely in proteins. The most prominent are the $\alpha$ helix and $\beta$ conformations described below. Using fundamental chemical principles and a few experimental observations, Linus Pauling and Robert Corey predicted the existence of these secondary structures in 1951, several years before the first complete protein structure was elucidated.

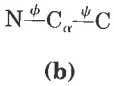
In considering secondary structure, it is useful to classify proteins into two major groups: fibrous proteins, having polypeptide chains arranged in long strands or sheets, and globular proteins, with polypeptide chains folded into a spherical or globular shape. Fibrous proteins play important structural roles in the anatomy and physiology of vertebrates, providing external protection, support, shape, and form. They may constitute one-half or more of the total body protein in larger animals. Most enzymes and peptide hormones are globular proteins. Globular proteins tend to be structurally complex, often containing several types of secondary structure; fibrous proteins usually consist largely of a single type of secondary structure. Because of this structural simplicity, certain fibrous proteins played a key role in the development of the modern understanding of protein structure and provide particularly clear examples of the relationship between structure and function; they are considered in some detail after the general discussion of secondary structure.

## The Peptide Bond Is Rigid and Planar

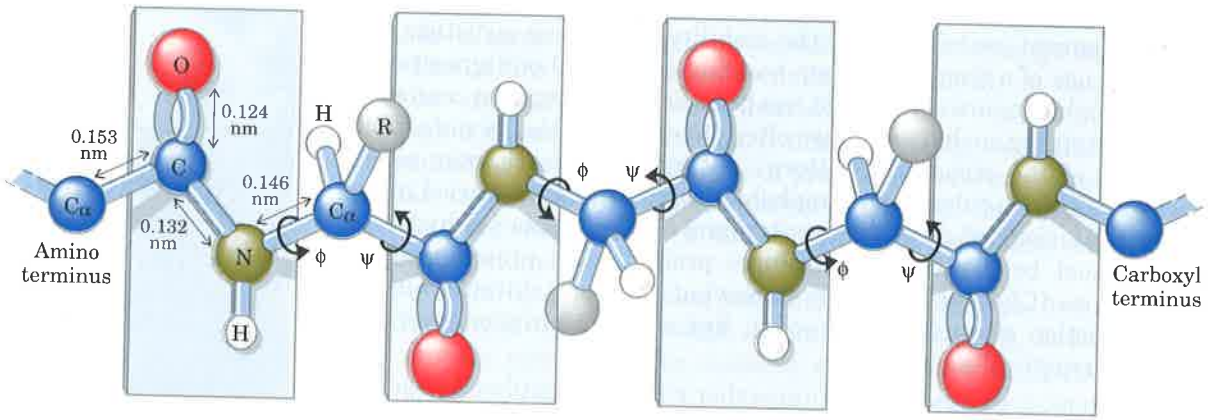
Pauling and Corey began their work on protein structure in the late 1930s by first focusing on the structure of the peptide bond. The $\alpha$ carbons of adjacent amino acids are separated by three covalent bonds, arranged $C_\alpha$—C—N—$C_\alpha$. X-ray diffraction studies of crystals of amino acids and of simple dipeptides and tripeptides demonstrated that the amide C—N bond in a peptide is somewhat shorter than the C—N bond in a simple amine and that the atoms associated with the bond are coplanar. This indicated a resonance or partial sharing of two pairs of electrons between the carbonyl oxygen and the amide nitrogen (Fig.



(a)



(b)




(c)

**Figure 7–4** **(a)** The planar peptide group. Each peptide bond has some double-bond character due to resonance and cannot rotate. The carbonyl oxygen has a partial negative charge and the amide nitrogen a partial positive charge, setting up a small electric dipole. Note that the oxygen and hydrogen atoms in the plane are on opposite sides of the C—N bond. This is the trans configuration. Virtually all peptide bonds in proteins occur in this configuration, although an exception is noted in Fig. 7–10. **(b)** Three bonds separate sequential $C_\alpha$ carbons in a polypeptide chain. The N—$C_\alpha$ and $C_\alpha$—C bonds can rotate, with bond angles designated $\phi$ and $\psi$, respectively. **(c)** Limited rotation can occur around two of the three types of bonds in a polypeptide chain. The C—N bonds in the planar peptide groups (shaded in blue), which make up one-third of all the backbone bonds, are not free to

7–4a). The oxygen has a partial negative charge and the nitrogen a partial positive charge, setting up a small electric dipole. The four atoms of the peptide group lie in a single plane, in such a way that the oxygen atom of the carbonyl group and the hydrogen atom of the amide nitrogen are trans to each other. From these studies Pauling and Corey concluded that the amide C—N bonds are unable to rotate freely because of their partial double-bond character. The backbone of a polypeptide chain can thus be pictured as a series of rigid planes separated by substituted methylene groups, —CH(R)— (Fig. 7–4c). The rigid peptide bonds limit the number of conformations that can be assumed by a polypeptide chain.

Rotation is permitted about the N—$C_\alpha$ and the $C_\alpha$—C bonds. By convention the bond angles resulting from rotations are labeled $\phi$ (phi) for the N—$C_\alpha$ bond and $\psi$ (psi) for the $C_\alpha$—C bond. Again by convention, both $\phi$ and $\psi$ are defined as 0° in the conformation in which the two peptide bonds connected to a single $\alpha$ carbon are in the same plane, as shown in Figure 7–4d. In principle, $\phi$ and $\psi$ can have any value between −180° and +180°, but many values of $\phi$ and $\psi$ are prohibited by steric interference between atoms in the polypeptide backbone and amino acid side chains. The conformation in which $\phi$ and $\psi$ are both 0° is prohibited for this reason; this is used merely as a reference point for describing the angles of rotation.

Every possible secondary structure is described completely by the two bond angles $\phi$ and $\psi$ that are repeated at each residue. Allowed values for $\phi$ and $\psi$ can be shown graphically by simply plotting $\psi$ versus $\phi$, an arrangement known as a **Ramachandran plot.** The Ramachandran plot in Figure 7–5 shows the conformations permitted for most amino acid residues.

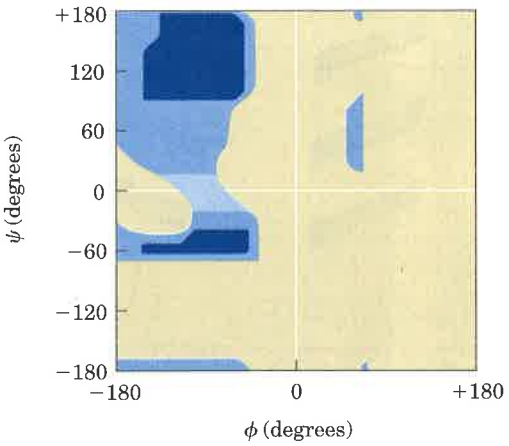
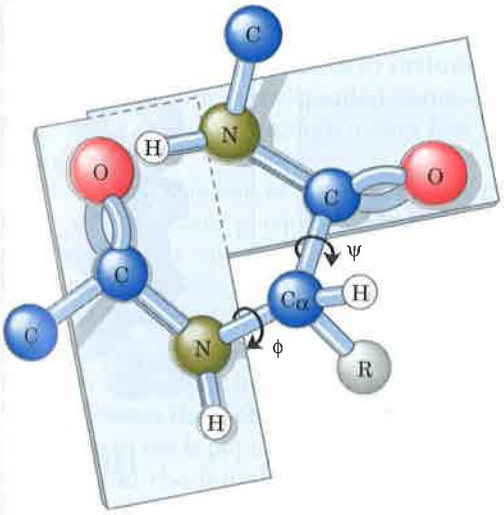



**Figure 7–5** A Ramachandran plot. The theoretically allowed conformations of peptides are shown, defined by the values of $\phi$ and $\psi$. The shaded areas reflect conformations that can be take up by all amino acids (dark shading) or all except valine and isoleucine (medium shading); the lightest shading reflects conformations that are somewhat unstable but are found in some protein structures.




**(d)**

rotate. Other single bonds in the backbone may also be rotationally hindered, depending on the size and charge of the R groups. **(d)** By convention, $\phi$ and $\psi$ are both defined as 0° when the two peptide bonds flanking an $\alpha$ carbon are in the same plane. In a protein, this conformation is prohibited by steric overlap between a carbonyl oxygen and an $\alpha$-amino hydrogen atom.

## The α Helix Is a Common Protein Secondary Structure

Pauling and Corey were aware of the importance of hydrogen bonds in orienting polar chemical groups such as the —C=O and —N—H groups of the peptide bond. They also had the experimental results of William Astbury, who in the 1930s had conducted pioneering x-ray studies of proteins. Astbury demonstrated that the protein that makes up hair and wool (the fibrous protein α-keratin) has a regular structure that repeats every 0.54 nm. With this information and their data on the peptide bond, and with the help of precisely constructed models, Pauling and Corey set out to determine the likely conformations of protein molecules.

The simplest arrangement the polypeptide chain could assume with its rigid peptide bonds (but with the other single bonds free to rotate) is a helical structure, which Pauling and Corey called the **α helix** (Fig. 7–6). In this structure the polypeptide backbone is tightly wound around the long axis of the molecule, and the R groups of the amino acid residues protrude outward from the helical backbone. The repeating unit is a single turn of the helix, which extends about 0.56 nm along the long axis, corresponding closely to the periodicity

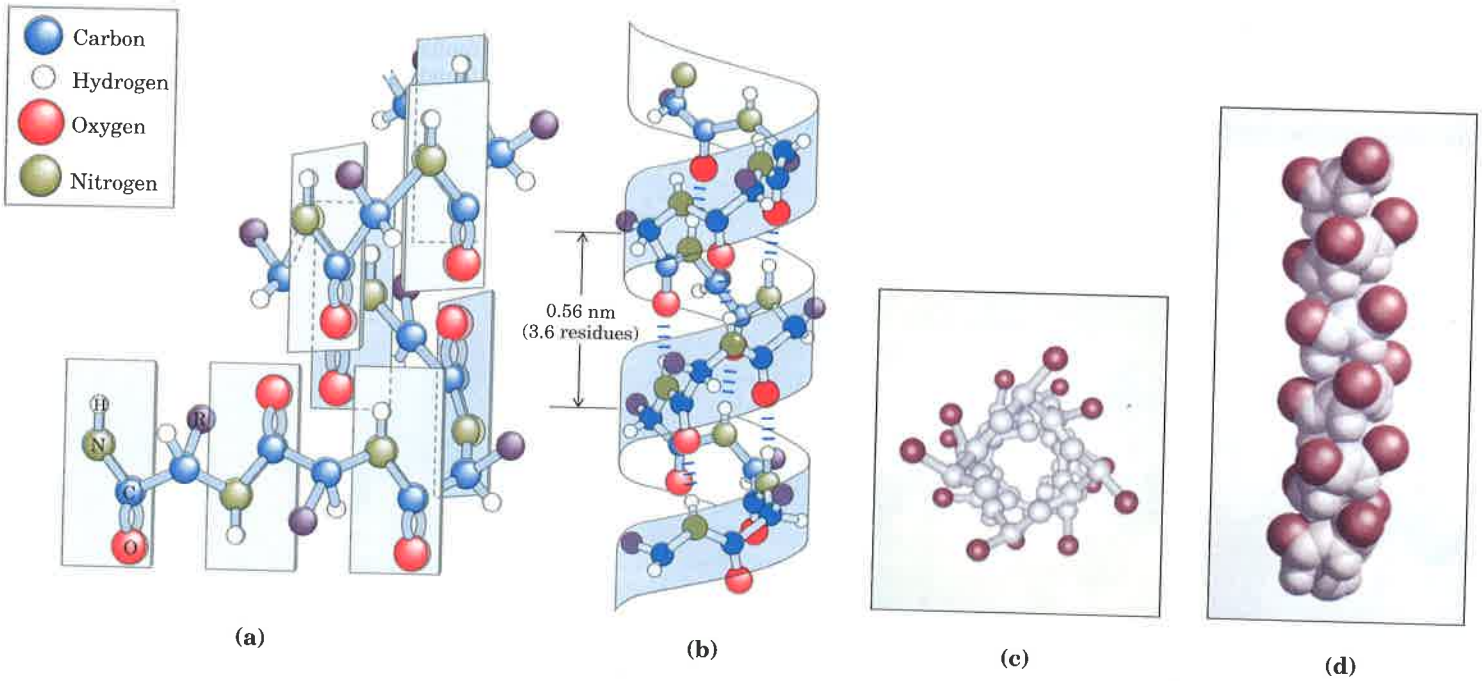
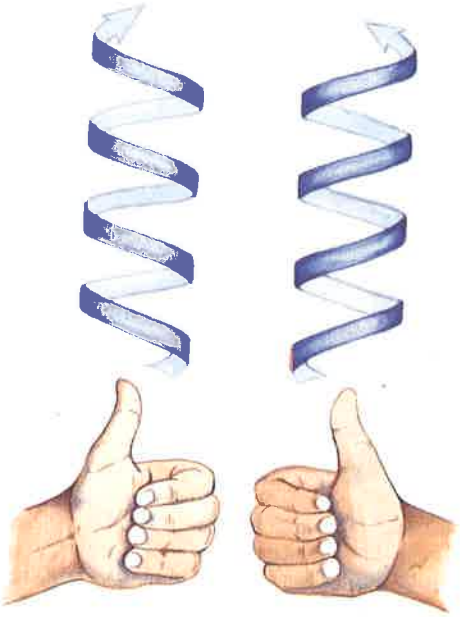



**Figure 7–6** Four models of the α helix, showing different aspects of its structure. **(a)** Formation of a right-handed α helix. The planes of the rigid peptide bonds are parallel to the long axis of the helix. **(b)** Ball-and-stick model of a right-handed α helix, showing the intrachain hydrogen bonds. The repeat unit is a single turn of the helix, 3.6 residues. **(c)** The α helix as viewed from one end, looking down the longitudinal axis. Note the positions of the R groups, represented by red spheres. **(d)** A space-filling model of the α helix.

## BOX 7–1 Knowing the Right Hand from the Left

There is a simple method for determining the handedness of a helical structure, whether right-handed or left-handed. Make fists of your two hands with thumbs outstretched and pointing away from you. Looking at your right hand, think of a helix spiraling away in the direction indicated by your right thumb, and the spiral occurring in the direction in which the other four fingers are curled as shown (clockwise). The resulting helix is right-handed. Repeating the process with your left hand will produce an image of a left-handed helix, which rotates in the counterclockwise direction as it spirals away from you.



Astbury observed on x-ray analysis of hair keratin. The amino acid residues in an $\alpha$ helix have conformations with $\psi = -45°$ to $-50°$ and $\phi = -60°$, and each helical turn includes 3.6 amino acids. The twisting of the helix has a right-handed sense (Box 7–1) in the most common form of the $\alpha$ helix, although a very few left-handed variants have been observed.

The $\alpha$ helix is one of two prominent types of secondary structure in proteins. It is the predominant structure in $\alpha$-keratins. In globular proteins, about one-fourth of all amino acid residues are found in $\alpha$ helices, the fraction varying greatly from one protein to the next.

Why does such a helix form more readily than many other possible conformations? The answer is, in part, that it makes optimal use of internal hydrogen bonds. The structure is stabilized by a hydrogen bond between the hydrogen atom attached to the electronegative nitrogen atom of each peptide linkage and the electronegative carbonyl oxygen atom of the fourth amino acid on the amino-terminal side of it in the helix (Fig. 7–6b). Every peptide bond of the chain participates in such hydrogen bonding. Each successive coil of the $\alpha$ helix is held to the adjacent coils by several hydrogen bonds, which in summation give the entire structure considerable stability.

Further model-building experiments have shown that an $\alpha$ helix can form with either L- or D-amino acids. However, all residues must be of one stereoisomeric series; a D-amino acid will disrupt a regular structure consisting of L-amino acids, and vice versa. Naturally occurring L-amino acids can form either right- or left-handed helices, but, with rare exceptions, only right-handed helices are found in proteins.

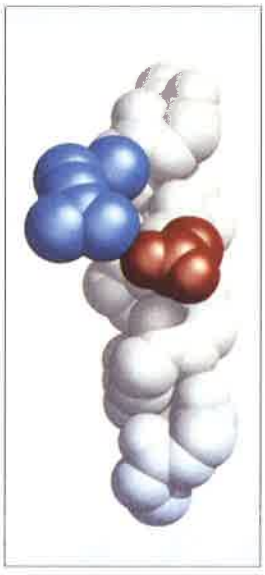


**Figure 7–7** Interactions between R groups of amino acids three residues apart in an $\alpha$ helix. An ionic interaction between $Asp^{100}$ and $Arg^{103}$ in an $\alpha$-helical region of the protein troponin C is shown in this space-filling model. The polypeptide backbone (carbons, $\alpha$-amino nitrogens, and $\alpha$-carbonyl oxygens) is shown in white for a helix segment about 12 amino acids long. The only side chains shown are the interacting Asp and Arg residues, with the aspartate in red and the arginine in blue. The side chain interaction illustrated occurs within the white connecting helix in Fig. 7–3.

## Amino Acid Sequence Affects $\alpha$ Helix Stability

Not all polypeptides can form a stable $\alpha$ helix. Additional interactions occur between amino acid side chains that can stabilize or destabilize this structure. For example, if a polypeptide chain has many Glu residues in a long block, this segment of the chain will not form an $\alpha$ helix at pH 7.0. The negatively charged carboxyl groups of adjacent Glu residues repel each other so strongly that they overcome the stabilizing influence of hydrogen bonds on the $\alpha$ helix. For the same reason, if there are many adjacent Lys and/or Arg residues, with positively charged R groups at pH 7.0, they will also repel each other and prevent formation of the $\alpha$ helix. The bulk and shape of certain R groups can also destabilize the $\alpha$ helix or prevent its formation. For example, Asn, Ser, Thr, and Leu residues tend to prevent formation of the $\alpha$ helix if they occur close together in the chain.

The twist of an $\alpha$ helix ensures that critical interactions occur between an amino acid side chain and the side chain three (and sometimes four) residues away on either side of it (Fig. 7–7). Positively charged amino acids are often found three residues away from negatively charged amino acids, permitting the formation of an ionic interaction. Two aromatic amino acids are often similarly spaced, resulting in a hydrophobic interaction.

A minor constraint on the formation of the $\alpha$ helix is the presence of Pro residues. In proline the nitrogen atom is part of a rigid ring (Fig. 5–6), and rotation about the N—$C_\alpha$ bond is not possible. In addition, the nitrogen atom of a Pro residue in peptide linkage has no substituent hydrogen-to-hydrogen bond with other residues. For these reasons, proline is only rarely found within an $\alpha$ helix.

A final factor affecting the stability of an $\alpha$ helix is the identity of the amino acids located near the ends of the $\alpha$-helical segment of a polypeptide. A small electric dipole exists in each peptide bond (see Fig. 7–4). These dipoles add across the hydrogen bonds in the helix so that the net dipole increases as helix length increases (Fig. 7–8). The four amino acids at either end of the helix do not participate fully in the helix hydrogen bonds. The partial positive and negative charges of the helix dipole actually reside on the peptide amino and carbonyl groups near the amino-terminal and carboxyl-terminal ends of the helix, respectively. For this reason, negatively charged amino acids are often found near the amino terminus of the helical segment, where they have a stabilizing interaction with the positive charge of the helix dipole; a positively charged amino acid at the amino-terminal end is destabilizing. The opposite is true at the carboxyl-terminal end of the helical segment.

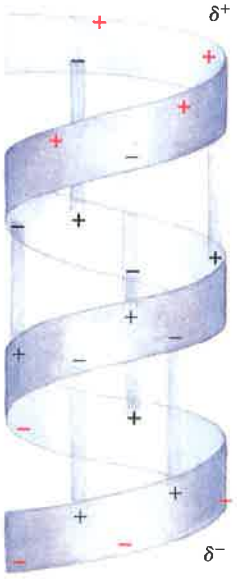


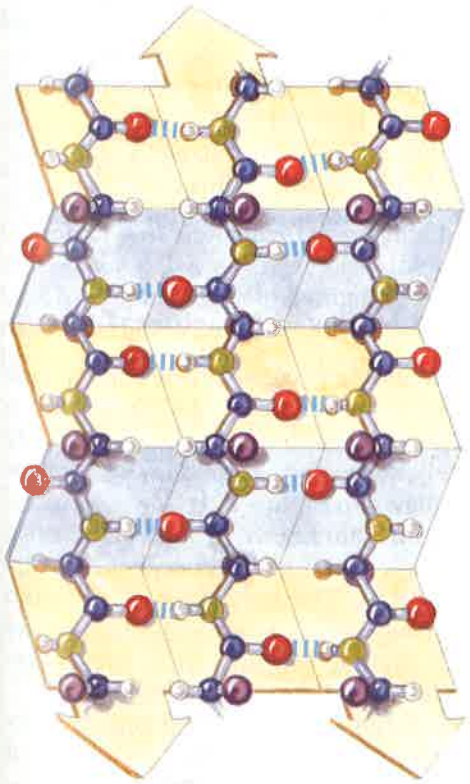

**Figure 7–8** The electric dipole of a peptide bond (Fig. 7–4a) is transmitted along an $\alpha$-helical segment through the intrachain hydrogen bonds, resulting in an overall helix dipole. In this illustration, the amino and carbonyl constituents of each peptide bond are indicated by + and − symbols, respectively. Unbonded amino and carbonyl constituents in the peptide bonds near either end of the $\alpha$-helical region are shown in red.

Thus there are five different kinds of constraints that affect the stability of an $\alpha$ helix: (1) the electrostatic repulsion (or attraction) between amino acid residues with charged R groups, (2) the bulkiness of adjacent R groups, (3) the interactions between amino acid side chains spaced three (or four) residues apart, (4) the occurrence of Pro residues, and (5) the interaction between amino acids at the ends of the helix and the electric dipole inherent to this structure.

## The $\beta$ Conformation Organizes Polypeptide Chains into Sheets

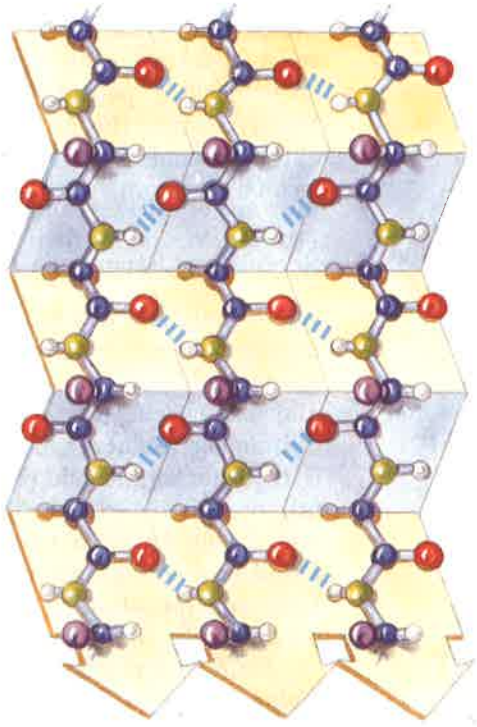
Pauling and Corey predicted a second type of repetitive structure, the **$\beta$ conformation.** This is the more extended conformation of the polypeptide chains, as seen in the silk protein fibroin (a member of a class of fibrous proteins called $\beta$-keratins), and its structure has been confirmed by x-ray analysis. In the $\beta$ conformation, which like the $\alpha$ helix is common in proteins, the backbone of the polypeptide chain is extended into a zigzag rather than helical structure (Fig. 7–9). In fibroin the zigzag polypeptide chains are arranged side by side to form a structure resembling a series of pleats; such a structure is called a $\beta$ pleated sheet. In the $\beta$ conformation the hydrogen bonds can be either intrachain, or interchain between the peptide linkages of adjacent polypeptide chains. All the peptide linkages of $\beta$-keratin participate in interchain hydrogen bonding. The R groups of adjacent amino acids protrude in opposite directions from the zigzag structure, creating an alternating pattern as seen in the side view (Fig. 7–9c).

**Figure 7–9** The $\beta$ conformation of polypeptide chains. Views show the R groups extending out from the $\beta$ pleated sheet and emphasize the pleated sheet described by the planes of the peptide bonds. Hydrogen-bond cross-links between adjacent chains are also shown. **(a)** Antiparallel $\beta$ sheets, in which the amino-terminal to carboxyl-terminal orientation of adjacent chains (arrows) is inverse. **(b)** Parallel $\beta$ sheets. **(c)** Silk fibers are made up of the protein fibroin. Its structure consists of layers of antiparallel $\beta$ sheets rich in Ala (purple) and Gly (yellow) residues. The small side chains interdigitate and allow close packing of each layered sheet, as shown in this side view.



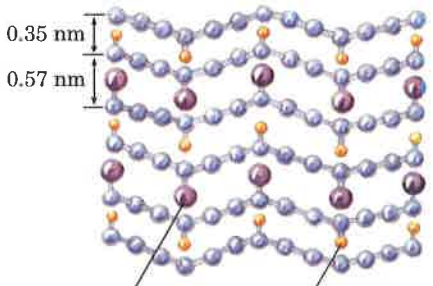
Antiparallel

(a)



Parallel

(b)




(c)

The adjacent polypeptide chains in a $\beta$ pleated sheet can be either parallel (having the same amino-to-carboxyl polypeptide orientation) or antiparallel (having the opposite amino-to-carboxyl orientation). The structures are similar, although the repeat period is shorter for the parallel conformation (0.65 nm, as opposed to 0.7 nm for antiparallel).

In some structural situations there are limitations to the kinds of amino acids that can occur in the $\beta$ structure. When two or more pleated sheets are layered closely together within a protein, the R groups of the amino acid residues on the contact surfaces must be relatively small. $\beta$-Keratins such as silk fibroin and the protein of spider webs have a very high content of Gly and Ala residues, those with the smallest R groups. Indeed, in silk fibroin Gly and Ala alternate over large parts of the sequence (Fig. 7–9c).

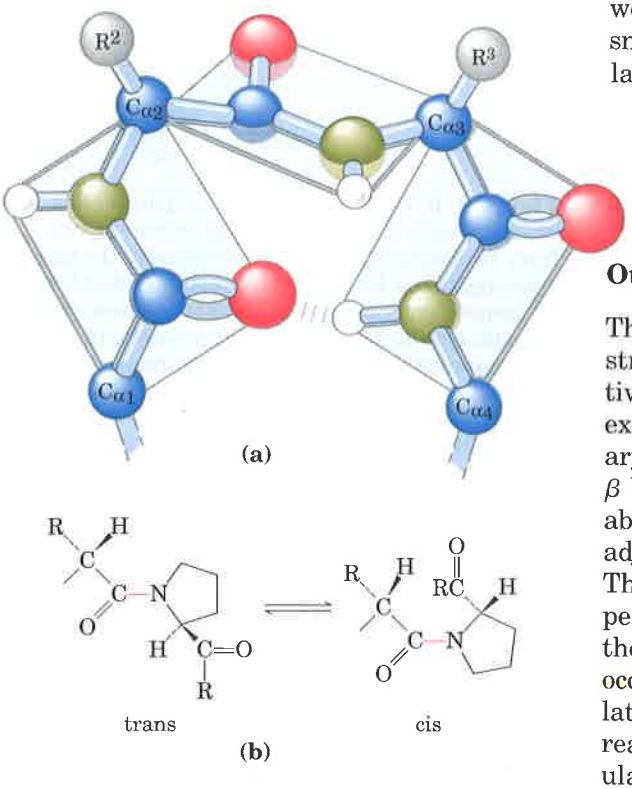



**Figure 7–10** Structure of a $\beta$ turn or $\beta$ bend. **(a)** Note the hydrogen bond between the peptide groups of the first and fourth residues involved in the bend. **(b)** The trans and cis isomers of a peptide bond involving the imino nitrogen of proline. Over 99.95% of the peptide bonds between amino acid residues other than Pro are in the trans configuration. About 6% of the peptide bonds involving the imino nitrogen of proline, however, are in the cis configuration, and many of these occur at $\beta$ turns.

## Other Secondary Structures Occur in Some Proteins

The $\alpha$ helix and the $\beta$ conformation are the major repetitive secondary structures easily recognized in a wide variety of proteins. Other repetitive structures exist, often in only one or a few specialized proteins. An example is the collagen helix (see Fig. 7–14). One other type of secondary structure is common enough to deserve special mention. This is a $\beta$ bend or $\beta$ turn (Fig. 7–10), often found where a polypeptide chain abruptly reverses direction. (These turns often connect the ends of two adjacent segments of an antiparallel $\beta$ pleated sheet, hence the name.) The structure is a tight turn (~180°) involving four amino acids. The peptide groups flanking the first amino acid are hydrogen bonded to the peptide groups flanking the fourth. Gly and Pro residues often occur in $\beta$ turns, the former because it is small and flexible; and the latter because peptide bonds involving the imino nitrogen of proline readily assume the cis configuration (Fig. 7–10b), a form that is particularly amenable to a tight turn. $\beta$ Turns are often found near the surface of a protein.

## Secondary Structure Is Affected by Several Factors

The $\alpha$ helix and $\beta$ conformation are stable because steric repulsion is minimized and hydrogen bonding is maximized. As shown by a Ramachandran plot, these structures fall within a range of sterically allowed structures that is relatively restricted. Values of $\phi$ and $\psi$ for common secondary structures are shown in Figure 7–11. Most values of $\phi$ and $\psi$ for amino acid residues, taken from known protein structures, fall into the expected regions, with high concentrations near the $\alpha$ helix and $\beta$ conformation values as expected. The only amino acid often found in a conformation outside these regions is glycine. Because its hydrogen side chain is small, a Gly residue can take up many conformations that are sterically forbidden for other amino acids.

Some amino acids are accommodated in the different types of secondary structures better than others. An overall summary is presented in Figure 7–12. Some biases, such as the presence of Pro and Gly residues in $\beta$ turns, can be explained readily; other evident biases are not understood.

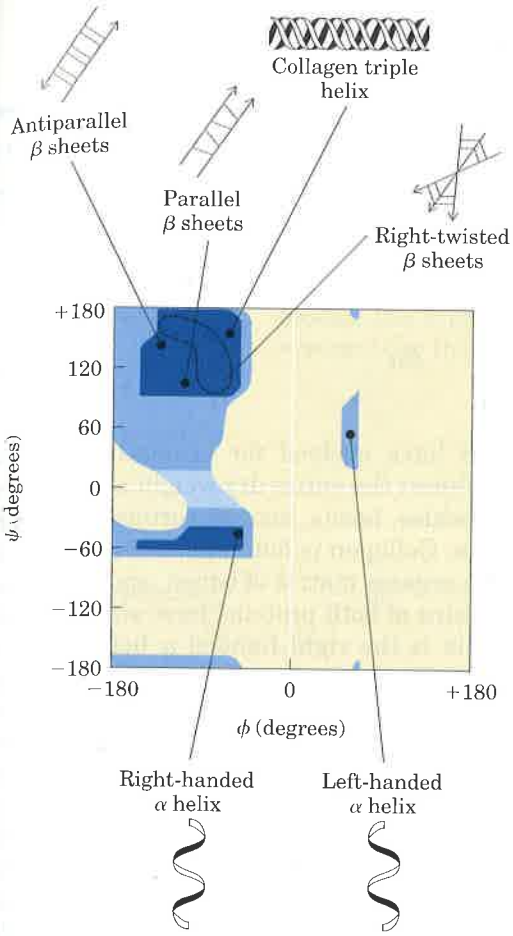



**Figure 7–11** A Ramachandran plot. The values of $\phi$ and $\psi$ for the various secondary structures are overlaid on the plot from Fig. 7–5.

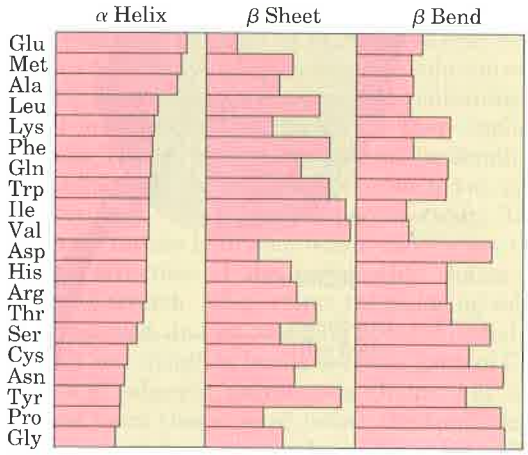



**Figure 7–12** Relative probabilities that a given amino acid will occur in the three common types of secondary structure.

## Fibrous Proteins Are Adapted for a Structural Function

α-Keratin, collagen, and elastin provide clear examples of the relationship between protein structure and biological function (Table 7–1). These proteins share properties that give strength and/or elasticity to structures in which they occur. They have relatively simple structures, and all are insoluble in water, a property conferred by a high concentration of hydrophobic amino acids both in the interior of the protein and on the surface. These proteins represent an exception to the rule that hydrophobic groups must be buried. The hydrophobic core of the molecule therefore contributes less to structural stability, and covalent bonds assume an especially important role.

**Table 7–1 Secondary structures and properties of fibrous proteins**

| Structure | Characteristics | Examples of occurrence |
|---|---|---|
| α Helix, cross-linked by disulfide bonds | Tough, insoluble protective structures of varying hardness and flexibility | α-Keratin of hair, feathers, and nails |
| β Conformation | Soft, flexible filaments | Fibroin of silk |
| Collagen triple helix | High tensile strength, without stretch | Collagen of tendons, bone matrix |
| Elastin chains cross-linked by desmosine and lysinonorleucine | Two-way stretch with elasticity | Elastin of ligaments |

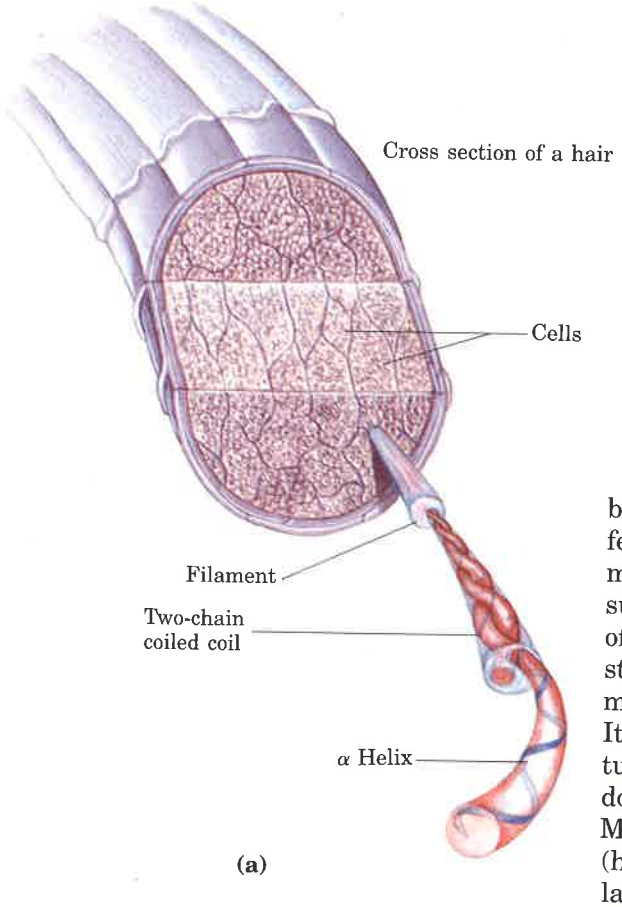



**Figure 7–13** **(a)** Hair α-keratin is an elongated α helix with somewhat thicker domains near the amino and carboxy termini. Pairs of these helices are interwound, probably in a left-handed sense, to form two-chain coiled coils. These then combine in higher-order structures called protofilaments and protofibrils, as shown in **(b)**. (About four protofibrils combine to form a filament.) The individual two-chain coiled coils in the various substructures also appear to be interwound, but the handedness of the interwinding and other structural details are unknown.

**α-Keratin** and **collagen** have evolved for strength. In vertebrates, α-keratins constitute almost the entire dry weight of hair, wool, feathers, nails, claws, quills, scales, horns, hooves, tortoise shell, and much of the outer layer of skin. Collagen is found in connective tissue such as tendons, cartilage, the organic matrix of bones, and the cornea of the eye. The polypeptide chains of both proteins have simple helical structures. The α-keratin helix is the right-handed α helix found in many other proteins (Fig. 7–13). However, the collagen helix is unique. It is left-handed (see Box 7–1) and has three amino acid residues per turn (Fig. 7–14). In both α-keratin and collagen, a few amino acids predominate. α-Keratin is rich in the hydrophobic residues Phe, Ile, Val, Met, and Ala. Collagen is 35% Gly, 11% Ala, and 21% Pro and Hyp (hydroxyproline; see Fig. 5–8). The unusual amino acid content of collagen is imposed by structural constraints unique to the collagen helix. The amino acid sequence in collagen is generally a repeating tripeptide unit, Gly–X–Pro or Gly–X–Hyp, where X can be any amino acid. The food product gelatin is derived from collagen. Although it is protein, it has little nutritional value because collagen lacks significant amounts of many amino acids that are essential in the human diet.

In both α-keratin and collagen, strength is amplified by wrapping multiple helical strands together in a superhelix, much the way strings are twisted to make a strong rope (Figs. 7–13, 7–14). In both proteins the helical path of the supertwists is opposite in sense to the twisting of the individual polypeptide helices, a conformation that permits the closest possible packing of the multiple polypeptide chains. The super-

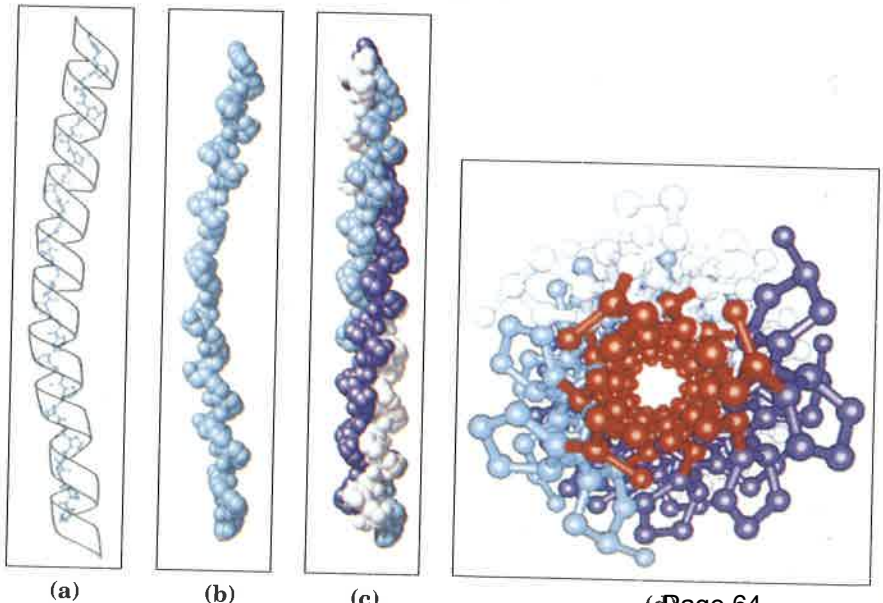
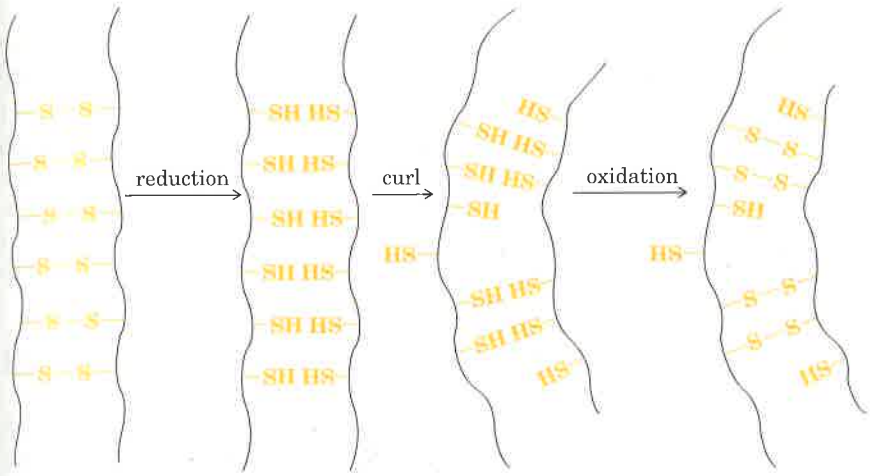



**Figure 7–14** Structure of collagen. The collagen helix is a repeating secondary structure unique to this protein. **(a)** The repeating tripeptide sequence Gly–X–Pro or Gly–X–Hyp adopts a left-handed helical structure with three residues per turn. The repeating sequence used to generate this model is Gly–Pro–Hyp. **(b)** Space-filling model of the collagen helix shown in **(a)**. **(c)** Three of these helices wrap around one another with a right-handed twist. The resulting three-stranded molecule is referred to as tropocollagen (see Fig. 7–15). **(d)** The three-stranded collagen superhelix shown from one end, in a ball-and-stick representation. Glycine residues are shown in red. Glycine, because of its small size, is required at the tight junction where the three chains are in contact.

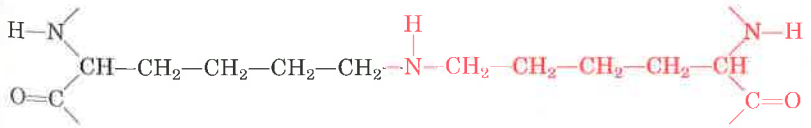
## BOX 7–2 Permanent Waving Is Biochemical Engineering




$\alpha$-Keratins exposed to moist heat can be stretched into the $\beta$ conformation, but on cooling revert to the $\alpha$-helical conformation spontaneously. This is because the R groups of $\alpha$-keratins are larger on average than those of $\beta$-keratins and thus are not compatible with a stable $\beta$ conformation. This characteristic of $\alpha$-keratins, as well as their content of disulfide cross-linkages, is the basis of permanent waving. The hair to be waved is first bent around a form of appropriate shape. A solution of a reducing agent, usually a compound containing a thiol or sulfhydryl group (—SH), is then applied with heat. The reducing agent cleaves the disulfide cross-linkages by reducing each cystine to two cysteine residues, one in each adjacent chain. The moist heat breaks hydrogen bonds and causes the $\alpha$-helical structure of the polypeptide chains to uncoil and stretch. After a time the reducing solution is removed, and an oxidizing agent is added to establish *new* disulfide bonds between pairs of Cys residues of adjacent polypeptide chains, but not the same pairs that existed before the treatment. On washing and cooling the hair, the polypeptide chains revert to their $\alpha$-helical conformation. The hair fibers now curl in the desired fashion because new disulfide cross-linkages have been formed where they will exert some torsion or twist on the bundles of $\alpha$-helical coils in the hair fibers.

helical twisting is probably left-handed in $\alpha$-keratin (Fig. 7–13) and right-handed in collagen (Fig. 7–14). The tight wrapping of the collagen triple helix provides great tensile strength with no capacity to stretch: Collagen fibers can support up to 10,000 times their own weight and are said to have greater tensile strength than a steel wire of equal cross section.

The strength of these structures is also enhanced by covalent cross-links between polypeptide chains within the multi-helical "ropes" and between adjacent ones. In $\alpha$-keratin, the cross-links are contributed by disulfide bonds (Box 7–2). In the hardest and toughest $\alpha$-keratins, such as those of tortoise shells and rhinoceros horns, up to 18% of the residues are cysteines involved in disulfide bonds. The arrangement of $\alpha$-keratin to form a hair fiber is shown in Figure 7–13. In collagen, the cross-links are contributed by an unusual type of covalent link between two Lys residues that creates a nonstandard amino acid residue called lysinonorleucine, found only in certain fibrous proteins.



Polypeptide chain — Lys residue minus $\epsilon$-amino group (norleucine) — Lys residue — Polypeptide chain

Lysinonorleucine

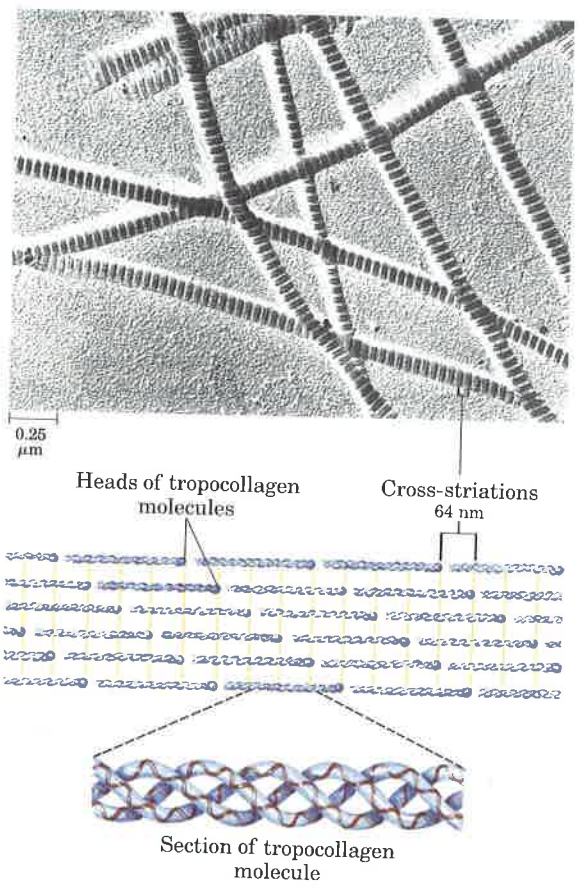



**Figure 7–15** The structure of collagen fibers. Tropocollagen ($M_r$ 300,000) is a rod-shaped molecule, about 300 nm long and only 1.5 nm thick. The three helically intertwined polypeptides are of equal length, each having about 1,000 amino acid residues. In some collagens all three chains are identical in amino acid sequence, but in others two chains are identical and the third differs. The heads of adjacent molecules are staggered, and the alignment of the head groups of every fourth molecule produces characteristic cross-striations 64 nm apart that are evident in an electron micrograph.

Collagen fibrils consist of recurring three-stranded polypeptide units called tropocollagen, arranged head to tail in parallel bundles (Fig. 7–15). The rigid, brittle character of the connective tissue in older people is the result of an accumulation of covalent cross-links in collagen as we age.

Human genetic defects involving collagen illustrate the close relationship between amino acid sequence and three-dimensional structure in this protein. Osteogenesis imperfecta results in abnormal bone formation in human babies. Ehlers-Danlos syndrome is characterized by loose joints. Both can be lethal and both result from the substitution of a Cys or Ser residue, respectively, for a Gly (a different Gly residue in each case) in the amino acid sequence of collagen. These seemingly small substitutions have a catastrophic effect on collagen function because they disrupt the Gly–X–Pro repeat that gives collagen its unique helical structure.

Elastic connective tissue contains the fibrous protein **elastin,** which resembles collagen in some of its properties but is very different in others. The polypeptide subunit of elastin fibrils is tropoelastin ($M_r$ 72,000), containing about 800 amino acid residues. Like collagen, it is rich in Gly and Ala residues. Tropoelastin differs from tropocollagen in having many Lys but few Pro residues; it forms a special type of helix, different from the $\alpha$ helix and the collagen helix. Tropoelastin consists of lengths of helix rich in Gly residues separated by short regions containing Lys and Ala residues. The helical portions stretch on applying tension but revert to their original length when tension is released.

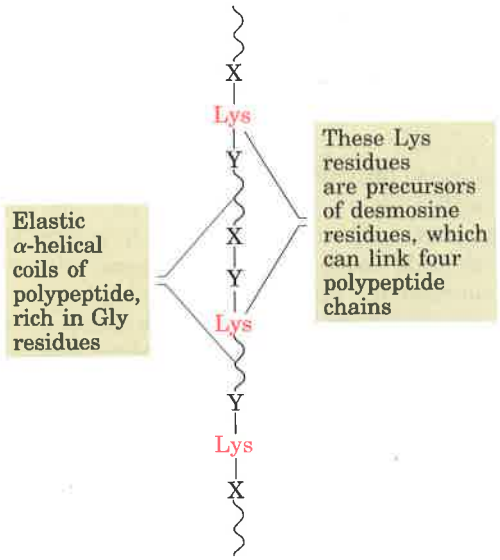
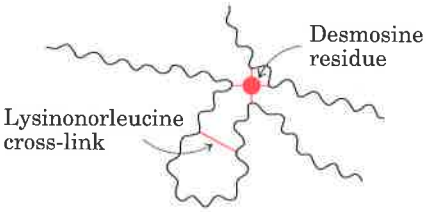
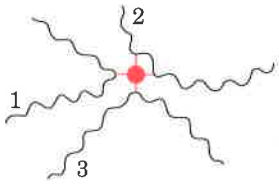



**Figure 7–16** Tropoelastin molecules and their linkage to form a network of polypeptide chains in elastin. Elastin consists of tropoelastin molecules cross-linked to give two-dimensional or three-dimensional elasticity. In addition to desmosine residues (in red), which can link two, three, or four tropoelastin molecules, as shown, elastin contains other kinds of cross-linkages, such as lysinonorleucine, also designated in red.

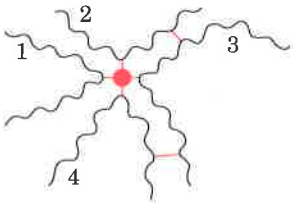
The regions containing Lys residues form covalent cross-links. Four Lys side chains come together and are enzymatically converted into desmosine (see Fig. 5–8) and a related compound, isodesmosine; these amino acids are found only in elastin. Lysinonorleucine (p. 173) also occurs in elastin. These nonstandard amino acids are capable of joining tropoelastin chains into arrays that can be stretched reversibly in all directions (Fig. 7–16).




Cross-linking two chains




Linking three chains




Linking four chains

## Protein Tertiary Structure

Although fibrous proteins generally have only one type of secondary structure, globular proteins can incorporate several types of secondary structure in the same molecule. Globular proteins—including enzymes, transport proteins, some peptide hormones, and immunoglobulins—are folded structures much more compact than $\alpha$ or $\beta$ conformations (as shown for serum albumin in Figure 7–17).

The three-dimensional arrangement of all atoms in a protein is referred to as the tertiary structure, and this now becomes our focus. Whereas the secondary structure of polypeptide chains is determined by the *short-range* structural relationship of amino acid residues, tertiary structure is conferred by *longer-range* aspects of amino acid sequence. Amino acids that are far apart in the polypeptide sequence and reside in different types of secondary structure may interact when the

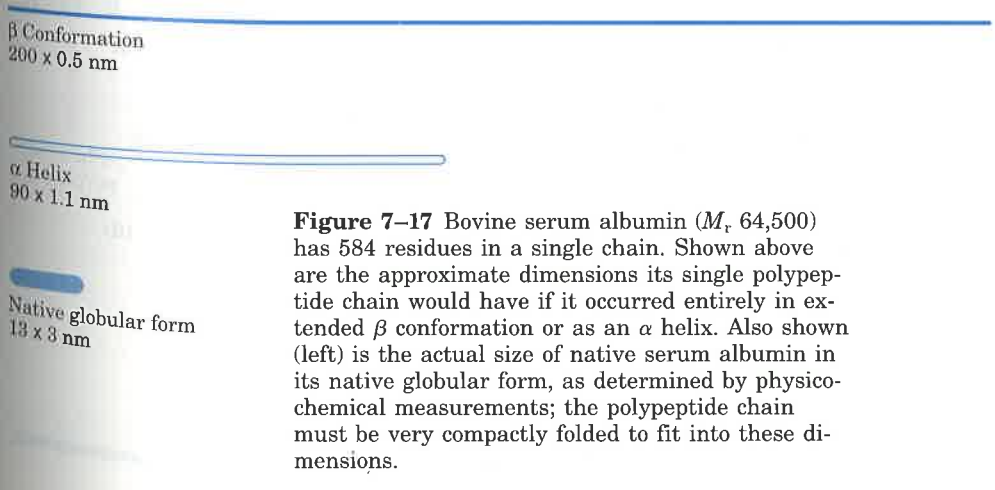



**Figure 7–17** Bovine serum albumin ($M_r$ 64,500) has 584 residues in a single chain. Shown above are the approximate dimensions its single polypeptide chain would have if it occurred entirely in extended $\beta$ conformation or as an $\alpha$ helix. Also shown (left) is the actual size of native serum albumin in its native globular form, as determined by physicochemical measurements; the polypeptide chain must be very compactly folded to fit into these dimensions.

protein is folded. The formation of bends in the polypeptide chain during folding and the direction and angle of these bends are determined by the number and location of specific bend-producing amino acids, such as Pro, Thr, Ser, and Gly residues. Moreover, loops of the highly folded polypeptide chain are held in their characteristic tertiary positions by different kinds of weak-bonding interactions (and sometimes by covalent bonds such as disulfide cross-links) between R groups of adjacent loops.

We will now consider how secondary structures contribute to the tertiary folding of a polypeptide chain in a globular protein, and how this structure is stabilized by weak interactions, in particular by hydrophobic interactions involving nonpolar amino acid side chains in the tightly packed core of the protein.

## X-Ray Analysis of Myoglobin Revealed Its Tertiary Structure

The breakthrough in understanding globular protein structure came from x-ray diffraction studies of the protein myoglobin carried out by John Kendrew and his colleagues in the 1950s (Box 7–3). Myoglobin is a relatively small ($M_r$ 16,700), oxygen-binding protein of muscle cells that functions in the storage and transport of oxygen for mitochondrial oxidation of cell nutrients. Myoglobin contains a single polypeptide chain of 153 amino acid residues of known sequence and a single iron-porphyrin, or **heme,** group (Fig. 7–18), identical to that of hemoglobin, the oxygen-binding protein of erythrocytes. The heme group is responsible for the deep red-brown color of both myoglobin and hemoglobin. Myoglobin is particularly abundant in the muscles of diving mammals such as the whale, seal, and porpoise, whose muscles are so rich in this protein that they are brown. Storage of oxygen by muscle myoglobin permits these animals to remain submerged for long periods of time.

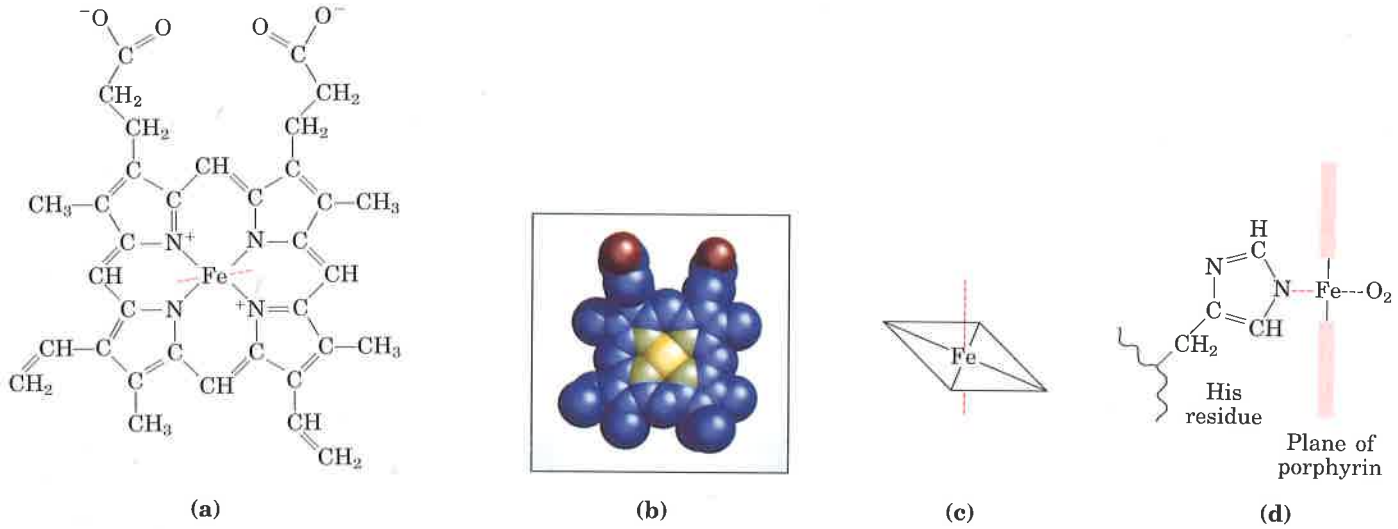



**Figure 7–18** The heme group, present in myoglobin, hemoglobin, cytochrome *b*, and many other heme proteins, consists of a complex organic ring structure, protoporphyrin, to which is bound an iron atom in its ferrous ($Fe^{2+}$) state. Two representations are shown in **(a)** and **(b). (c)** The iron atom has six coordination bonds, four in the plane of, and bonded to, the flat porphyrin molecule and two perpendicular to it. **(d)** In myoglobin and hemoglobin, one of the perpendicular coordination bonds is bound to a nitrogen atom of a His residue. The other is "open" and serves as the binding site for an $O_2$ molecule, as shown here in the edge view.

BOX 7–3

## X-Ray Diffraction

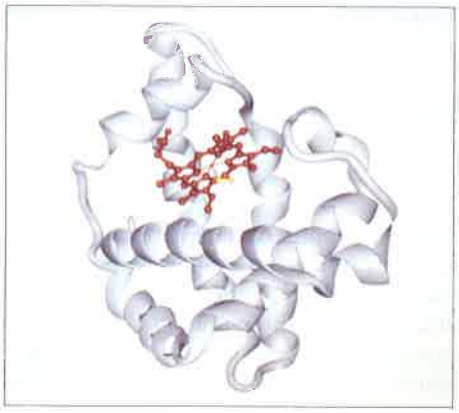
The spacing of atoms in a crystal lattice can be determined by measuring the angles and the intensities at which a beam of x rays of a given wavelength is diffracted by the electron shells around the atoms. For example, x-ray analysis of sodium chloride crystals shows that $Na^+$ and $Cl^-$ ions are arranged in a simple cubic lattice. The spacing of the different kinds of atoms in complex organic molecules, even very large ones such as proteins, can also be analyzed by x-ray diffraction methods. However, this is far more difficult than for simple salt crystals because the very large number of atoms in a protein molecule yields thousands of diffraction spots that must be analyzed by computer.

The process may be understood at an elementary level by considering how images are generated in a light microscope. Light from a point source is focused on an object. The light waves are scattered by the object, and these scattered waves are recombined by a series of lenses to generate an enlarged image of the object. The limit to the size of an object whose structure can be determined by such a system (i.e., its resolving power) is determined by the wavelength of the light. Objects smaller than half the wavelength of the incident light cannot be resolved. This is why x rays, with wavelengths in the range of a few tenths of a nanometer (often measured in angstroms, Å; 1 Å=0.1 nm), must be used for proteins. There are no lenses that can recombine x rays to form an image; the pattern of diffracted light is collected directly and converted into an image by computer analysis.

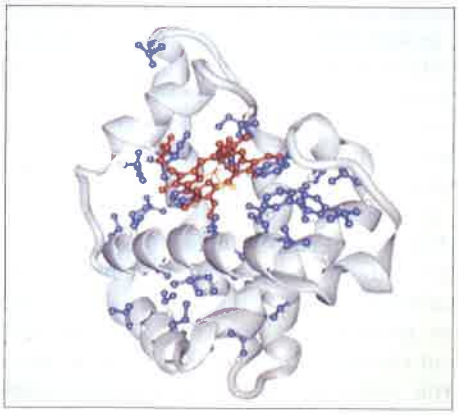
Operationally, there are several steps in x-ray structural analysis. The amount of information obtained depends on the degree of structural order in the sample. Some important structural parameters were obtained from early studies of the diffraction patterns of the fibrous proteins that occur in fairly regular arrays in hair and wool. More detailed three-dimensional structural information, however, requires a highly ordered crystal of a protein. Protein crystallization is something of an empirical science, and the structures of many important proteins are not yet known simply because they have proven difficult to crystallize. Once a crystal is obtained, it is placed in an x-ray beam between the x-ray source and a detector. A regular array of spots called reflections (Fig. 1) is generated by precessional motion of the crystal. The spots represent reflections of the x-ray beam, and each atom in a molecule makes a contribution to each spot. The overall pattern of spots is related to the structure of the protein through a mathematical device called a Fourier transform. The intensity of each spot is measured from the positions and intensities of the spots in several of these diffraction patterns, and the precise three-dimensional structure of the protein is calculated.

John Kendrew found that the x-ray diffraction pattern of crystalline myoglobin from muscles of the sperm whale is very complex, with nearly 25,000 reflections. Computer analysis of these reflections took place in stages. The resolution improved at each stage, until in 1959 the positions of virtually all the atoms in the protein could be determined. The amino acid sequence deduced from the structure agreed with that obtained by chemical analysis. The structures of hundreds of proteins have since been determined to a similar level of resolution, many of them much more complex than myoglobin.

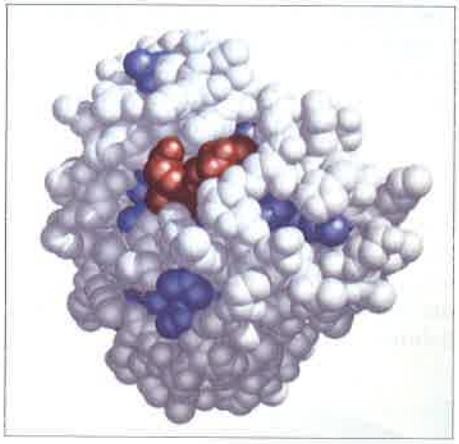
**Figure 1** Photograph of the x-ray diffraction pattern of crystalline sperm whale myoglobin.



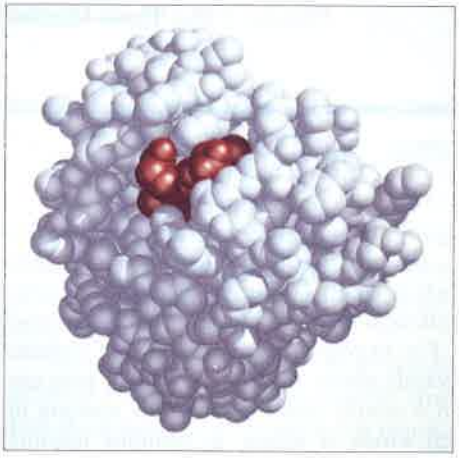
(a)



(c)



(d)



(b)

**Figure 7–19** Tertiary structure of sperm whale myoglobin. The orientation of the protein is the same in all panels; the heme group is shown in red. **(a)** The polypeptide backbone, shown in a ribbon representation of a type introduced by Jane Richardson; this highlights regions of secondary structure. The $\alpha$-helical regions in myoglobin are evident. Amino acid side chains are not shown. **(b)** A space-filling model, showing that the heme group is largely buried. All amino acid side chains are included. **(c)** A ribbon representation, including side chains (purple) for the hydrophobic residues Leu, Ile, Val, and Phe. **(d)** A space-filling model with all amino acid side chains. The hydrophobic residues are again shown in purple; most are not visible because they are buried in the interior of the protein.

Figure 7–19 shows several structural representations of myoglobin, illustrating how the polypeptide chain is folded in three dimensions—its tertiary structure. The backbone of the myoglobin molecule is made up of eight relatively straight segments of $\alpha$ helix interrupted by bends. The longest $\alpha$ helix has 23 amino acid residues and the shortest only seven; all are right-handed. More than 70% of the amino acids in the myoglobin molecule are in these $\alpha$-helical regions. X-ray analysis also revealed the precise position of each of the R groups, which occupy nearly all the open space between the folded loops.

Other important conclusions were drawn from the structure of myoglobin. The positioning of amino acid side chains reflects a structure that derives much of its stability from hydrophobic interactions. Most of the hydrophobic R groups are in the interior of the myoglobin molecule, hidden from exposure to water. All but two of the polar R groups are located on the outer surface of the molecule, and all of them are hydrated. The myoglobin molecule is so compact that in its interior there is room for only four molecules of water. This dense hydrophobic core is typical of globular proteins. The fraction of space occupied by atoms in an organic liquid is 0.25 to 0.35; in a typical solid the fraction is 0.75. In a protein the fraction is 0.72 to 0.76, very comparable to that in a solid. In this closely packed environment weak interactions strengthen and reinforce each other. For example, the nonpolar side chains in the core are so close together that short-range van der Waals interactions make a significant contribution to stabilizing hydrophobic interactions. By contrast, in an oil droplet suspended in water, the van der Waals interactions are minimal and the cohesiveness of the droplet is based almost exclusively on entropy.

The structure of myoglobin both confirmed some expectations and introduced some new elements of secondary structure. As predicted by Pauling and Corey, all the peptide bonds are in the planar trans configuration. The $\alpha$ helices in myoglobin provided the first direct experimental evidence for the existence of this type of secondary structure. Each of the four Pro residues of myoglobin occurs at a bend (recall that the rigid R group of proline is largely incompatible with $\alpha$-helical structure). Other bends contain Ser, Thr, and Asn residues, which are among the amino acids that tend to be incompatible with $\alpha$-helical structure if they are in close proximity (p. 168).

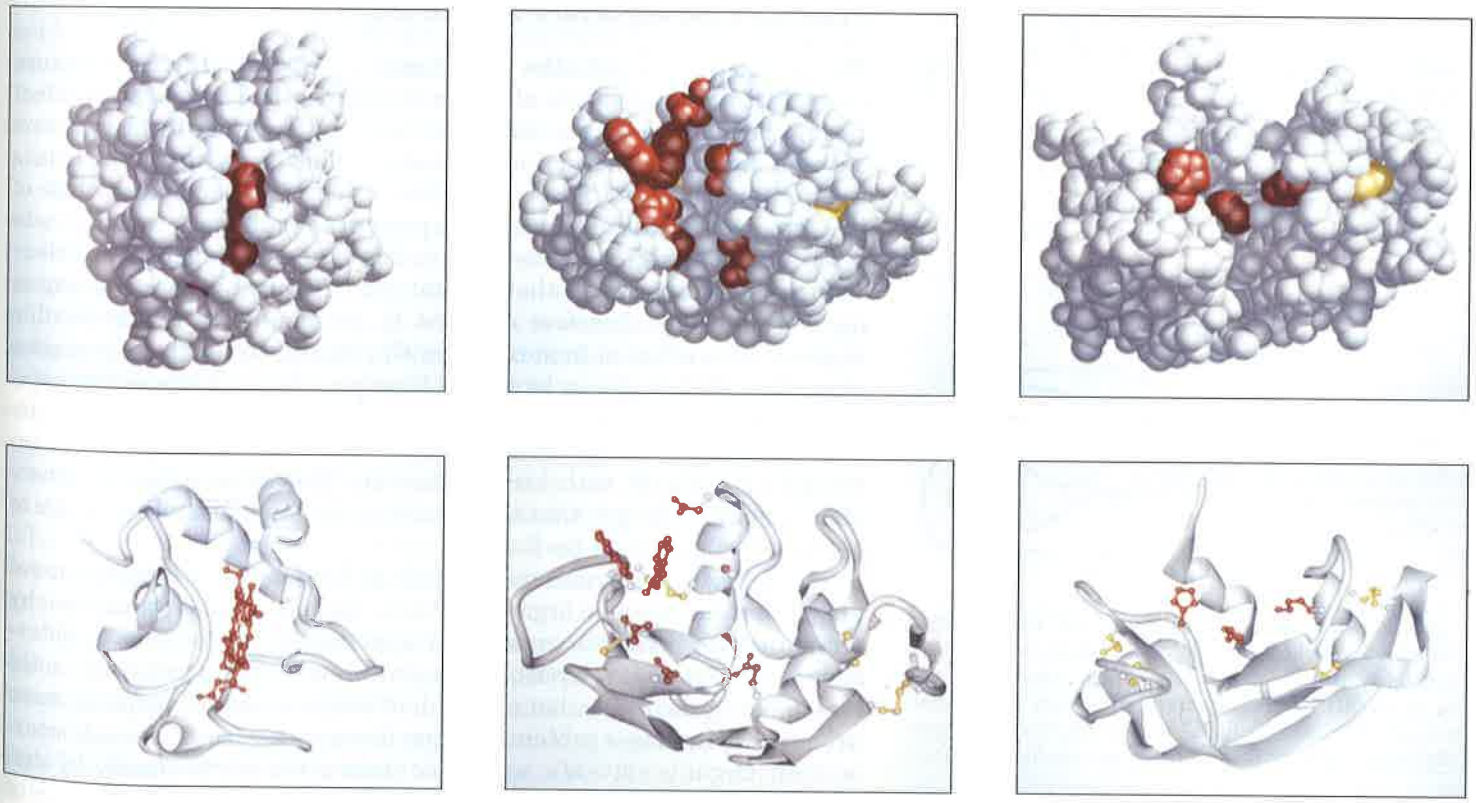
The flat heme group rests in a crevice, or pocket, in the myoglobin molecule. The iron atom in the center of the heme group has two bonding (coordination) positions perpendicular to the plane of the heme.

One of these is bound to the R group of the His residue at position 93; the other is the site to which an $O_2$ molecule is bound. Within this pocket, the accessibility of the heme group to solvent is highly restricted. This is important for function because free heme groups in an oxygenated solution are rapidly oxidized from the ferrous ($Fe^{2+}$) form, which is active in the reversible binding of $O_2$, to the ferric ($Fe^{3+}$) form, which does not bind $O_2$.

## Proteins Differ in Tertiary Structure

With the elucidation of the tertiary structures of hundreds of other globular proteins by x-ray analysis, it is clear that myoglobin represents only one of many ways in which a polypeptide chain can be folded. In Figure 7–20 the structures of cytochrome *c*, lysozyme, and ribonuclease are compared. All have different amino acid sequences and different tertiary structures, reflecting differences in function. Like myoglobin, cytochrome *c* is a small heme protein ($M_r$ 12,400) containing a single polypeptide chain of about 100 residues and a single heme group, which in this case is covalently attached to the polypeptide. It functions as a component of the respiratory chain of mitochondria (Chapter 18). X-ray analysis of cytochrome *c* (Fig. 7–20) shows that only about 40% of the polypeptide is in $\alpha$-helical segments, compared with almost 80% of the myoglobin chain. The rest of the cytochrome *c* chain contains bends, turns, and irregularly coiled and extended segments. Thus, cytochrome *c* and myoglobin differ markedly in structure, even though both are small heme proteins.

**Figure 7–20** The three-dimensional structures of three small proteins: cytochrome *c*, lysozyme, and ribonuclease. For lysozyme and ribonuclease the active site of the enzyme faces the viewer. Key functional groups (the heme in cytochrome *c*, and amino acid side chains in the active site of lysozyme and ribonuclease) are shown in red; disulfide bonds are shown in yellow. Two representations of each protein are shown: a space-filling model and a ribbon representation. In the ribbon depictions, the $\beta$ structures are represented by flat arrows and the $\alpha$ helices by spiral ribbons; the orientation in each case is the same as that of the space-filling model, to facilitate comparison.



**Cytochrome *c*** **Lysozyme** **Ribonuclease**

**Table 7–2 Approximate amounts of α helix and β conformation in some single-chain proteins***

| Protein (total residues) | Residues (%) | |
|---|---|---|
| | α Helix | β Conformation |
| Myoglobin (153) | 78 | 0 |
| Cytochrome *c* (104) | 39 | 0 |
| Lysozyme (129) | 40 | 12 |
| Ribonuclease (124) | 26 | 35 |
| Chymotrypsin (247) | 14 | 45 |
| Carboxypeptidase (307) | 38 | 17 |

*Source:* Data from Cantor, C.R. & Schimmel, P.R. (1980) *Biophysical Chemistry*, Part I: *The Conformation of Biological Macromolecules*, p. 100, W.H. Freeman and Company, New York.

* Portions of the polypeptide chains that are not accounted for by α helix or β conformation consist of bends and irregularly coiled or extended stretches. Segments of α helix and β conformation sometimes deviate slightly from their normal dimensions and geometry.

Lysozyme ($M_r$ 14,600) is an enzyme in egg white and human tears that catalyzes the hydrolytic cleavage of polysaccharides in the protective cell walls of some families of bacteria. Lysozyme is so named because it can lyse, or degrade, bacterial cell walls and thus serve as a bactericidal agent. Like cytochrome *c*, about 40% of its 129 amino acid residues are in α-helical segments, but the arrangement is different and some β structure is also present. Four disulfide bonds contribute stability to this structure. The α helices line a long crevice in the side of the molecule (Fig. 7–20), called the active site, which is the site of substrate binding and action. The bacterial polysaccharide that is the substrate for lysozyme fits into this crevice.

Ribonuclease, another small globular protein ($M_r$ 13,700), is an enzyme secreted by the pancreas into the small intestine, where it catalyzes the hydrolysis of certain bonds in the ribonucleic acids present in ingested food. Its tertiary structure, determined by x-ray analysis, shows that little of its 124 amino acid polypeptide chain is in α-helical conformation, but it contains many segments in the β conformation. Like lysozyme, ribonuclease has four disulfide bonds between loops of the polypeptide chain (Fig. 7–20).

Table 7–2 shows the relative percentages of α helix and β conformation among several small, single-chain, globular proteins. Each of these proteins has a distinct structure, adapted for its particular biological function. These proteins do share several important properties, however. Each is folded compactly, and in each case the hydrophobic amino acid side chains are oriented toward the interior (away from water) and the hydrophilic side chains are on the surface. These specific structures are also stabilized by a multitude of hydrogen bonds and some ionic interactions.

## Proteins Lose Structure and Function on Denaturation

The way to demonstrate the importance of a specific protein structure for biological function is to alter the structure and determine the effect on function. One extreme alteration is the total loss or randomization of three-dimensional structure, a process called **denaturation.** This is the familiar process that occurs when an egg is cooked. The white of the egg, which contains the soluble protein egg albumin, coagulates to a white solid on heating. It will not redissolve on cooling to yield a clear solution of protein as in the original unheated egg white. Heating of egg albumin has therefore changed it, seemingly in an irreversible manner. This effect of heat occurs with virtually all globular proteins, regardless of their size or biological function, although the precise temperature at which it occurs may vary and it is not always irreversible. The change in structure brought about by denaturation is almost invariably associated with loss of function. This is an expected consequence of the principle that the specific three-dimensional structure of a protein is critical to its function.

Proteins can be denatured not only by heat, but also by extremes of pH, by certain miscible organic solvents such as alcohol or acetone, by certain solutes such as urea, or by exposure of the protein to detergents. Each of these denaturing agents represents a relatively mild treatment in the sense that no covalent bonds in the polypeptide chain are broken. Boiling a protein solution disrupts a variety of weak interactions. Organic solvents, urea, and detergents act primarily by disrupting the hydrophobic interactions that make up the stable core of globular proteins; extremes of pH alter the net charge on the protein,

causing electrostatic repulsion and disruption of some hydrogen bonding. Remember that the native structure of most proteins is only marginally stable. It is not necessary to disrupt *all* of the stabilizing weak interactions to reduce the thermodynamic stability to a level that is insufficient to keep the protein conformation intact.

## Amino Acid Sequence Determines Tertiary Structure

The most important proof that the tertiary structure of a globular protein is determined by its amino acid sequence came from experiments showing that denaturation of some proteins is reversible. Some globular proteins denatured by heat, extremes of pH, or denaturing reagents will regain their native structure and their biological activity, a process called **renaturation,** if they are returned to conditions in which the native conformation is stable.

A classic example is the denaturation and renaturation of ribonuclease. Purified ribonuclease can be completely denatured by exposure to a concentrated urea solution in the presence of a reducing agent. The reducing agent cleaves the four disulfide bonds to yield eight Cys residues, and the urea disrupts the stabilizing hydrophobic interactions, thus freeing the entire polypeptide from its folded conformation. Under these conditions the enzyme loses its catalytic activity and undergoes complete unfolding to a randomly coiled form (Fig. 7–21). When the urea and the reducing agent are removed, the randomly coiled, denatured ribonuclease spontaneously refolds into its correct tertiary structure, with full restoration of its catalytic activity (Fig. 7–21). The refolding of ribonuclease is so accurate that the four intrachain disulfide bonds are reformed in the same positions in the renatured molecule as in the native ribonuclease. In theory, the eight Cys residues could have recombined at random to form up to four disulfide bonds in 105 different ways. This classic experiment, carried out by Christian Anfinsen in the 1950s, proves that the amino acid sequence of the polypeptide chain of proteins contains all the information required to fold the chain into its native, three-dimensional structure.

The study of homologous proteins has strengthened this conclusion. We have seen that in a series of homologous proteins, such as cytochrome *c,* from different species, the amino acid residues at certain positions in the sequence are invariant, whereas at other positions the amino acids may vary (see Fig. 6–15). This is also true for myoglobins isolated from different species of whales, from the seal, and from some terrestrial vertebrates. The similarity of the tertiary structures and amino acid sequences of myoglobins from different sources led to the conclusion that the amino acid sequence of myoglobin somehow must determine its three-dimensional folding pattern, an idea substantiated by the similar structures found by x-ray analysis of myoglobins from different species. Other sets of homologous proteins also show this relationship; in each case there are sequence homologies as well as similar tertiary structures.

Many of the invariant amino acid residues of homologous proteins appear to occur at critical points along the polypeptide chain. Some are found at or near bends in the chain, others at cross-linking points between loops in the tertiary structure, such as Cys residues involved in disulfide bonds. Still others occur at the catalytic sites of enzymes or at the binding sites for prosthetic groups, such as the heme group of cytochrome *c.*

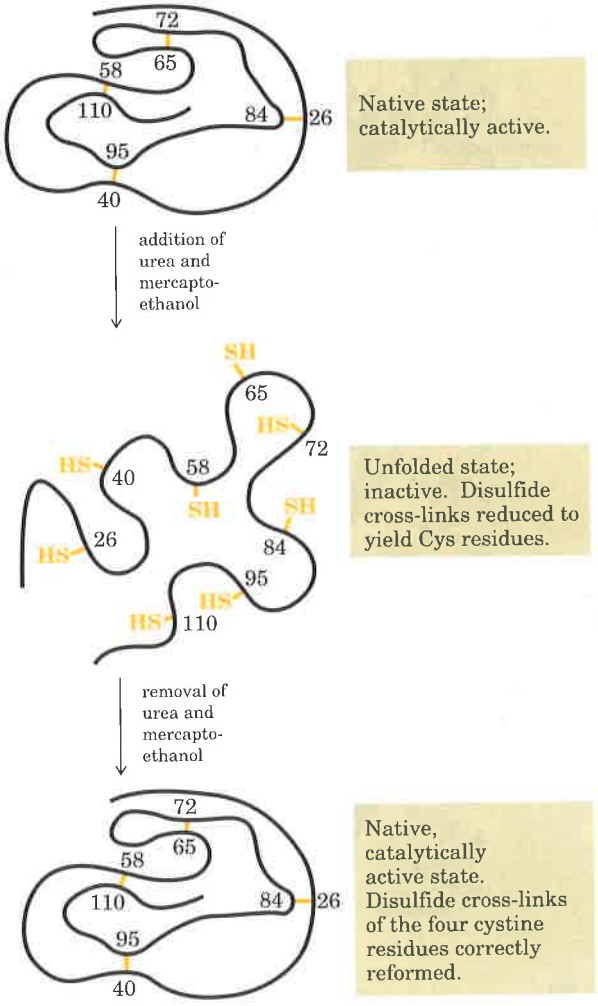



**Figure 7–21** Renaturation of unfolded, denatured ribonuclease, with reestablishment of correct disulfide cross-links. Urea is added to denature ribonuclease, and mercaptoethanol ($HOCH_2CH_2SH$) to reduce and thus cleave the disulfide bonds of the four cystine residues to yield eight cysteine residues.

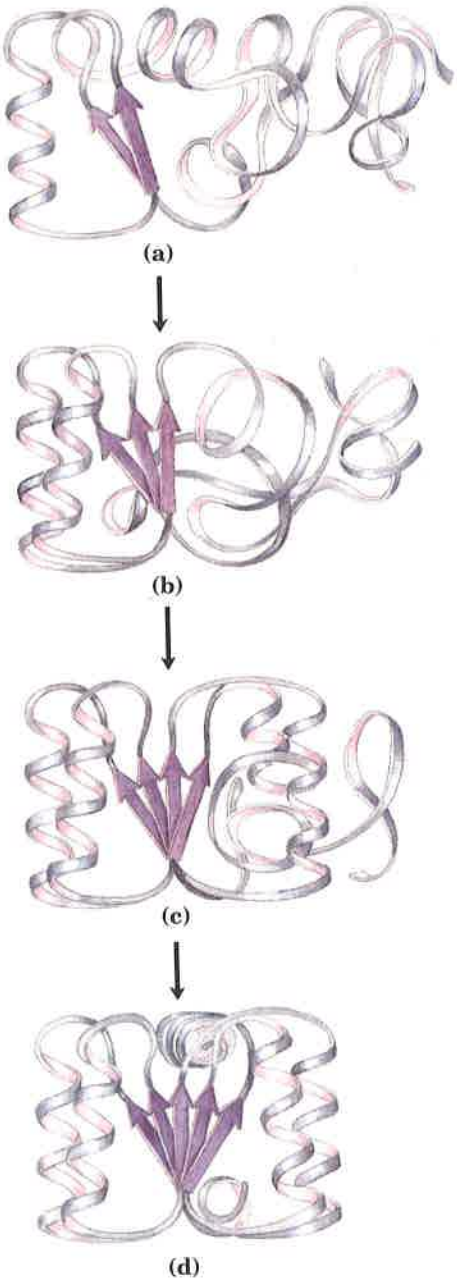



**Figure 7–22** A possible protein-folding pathway. **(a)** Protein folding often begins with spontaneous formation of a structural nucleus consisting of a few particularly stable regions of secondary structure. **(b)** As other regions adopt secondary structure, they are stabilized by long-range interactions with the structural nucleus. **(c)** The folding process continues until most of the polypeptide has assumed regular secondary structure. **(d)** The final structure generally represents the most thermodynamically stable conformation.

Looking at naturally occurring amino acid substitutions has an important limitation. Any change that abolishes the function of an essential protein (e.g., a change in an invariant residue) usually results in death of the organism very early in development. This severe form of natural selection eliminates many potentially informative changes from study. Fortunately, biochemists have devised methods to specifically alter amino acid sequences in the laboratory and examine the effects of these changes on protein structure and function. These methods are derived from recombinant DNA technology (Chapter 28) and rely on altering the genetic material encoding the protein. By this process, called **site-directed mutagenesis,** specific amino acid sequences can be changed by deleting, adding, rearranging, or substituting amino acid residues. The catalytic roles of certain amino acids lining the active sites of enzymes such as triose phosphate isomerase and chymotrypsin have been elucidated by substituting different amino acids in their place. The importance of certain amino acids in protein folding and structure is being addressed in the same way.

### Tertiary Structures Are Not Rigid

Although the native tertiary conformation of a globular protein is the thermodynamically most stable form its polypeptide chain can assume, this conformation must not be regarded as absolutely rigid. Globular proteins have a certain amount of flexibility in their backbones and undergo short-range internal fluctuations. Many globular proteins also undergo small conformational changes in the course of their biological function. In many instances, these changes are associated with the binding of a ligand. The term **ligand** in this context refers to a specific molecule that is bound by a protein (from Latin, *ligare,* "to tie" or "bind"). For example, the hemoglobin molecule, which we shall examine later in this chapter, has one conformation when oxygen is bound, and another when the oxygen is released. Many enzyme molecules also undergo a conformational change on binding their substrates, a process that is part of their catalytic action (Chapter 8).

### Polypeptides Fold Rapidly by a Stepwise Process

In living cells, proteins are made from amino acids at a very high rate. For example, *Escherichia coli* cells can make a complete, biologically active protein molecule containing 100 amino acid residues in about 5 s at 37 °C. Yet calculations show that at least $10^{50}$ yr would be required for a polypeptide chain of 100 amino acid residues to fold itself spontaneously by a random process in which it tries out all possible conformations around every single bond in its backbone until it finds its native, biologically active form. Thus protein folding cannot be a completely random, trial-and-error process. There simply must be shortcuts.

The folding pathway of a large polypeptide chain is unquestionably complicated, and the principles that guide this process have not yet been worked out in detail. For several proteins, however, there is evidence that folding proceeds through several discrete intermediates, and that some of the earliest steps involve local folding of regions of secondary structure. In one model (Fig. 7–22), the process is envisioned as hierarchical, following the levels of structure outlined at the beginning of this chapter. Local secondary structures would form first, followed by longer-range interactions between, say, two $\alpha$ helices with compatible amino acid side chains, a process continuing until folding

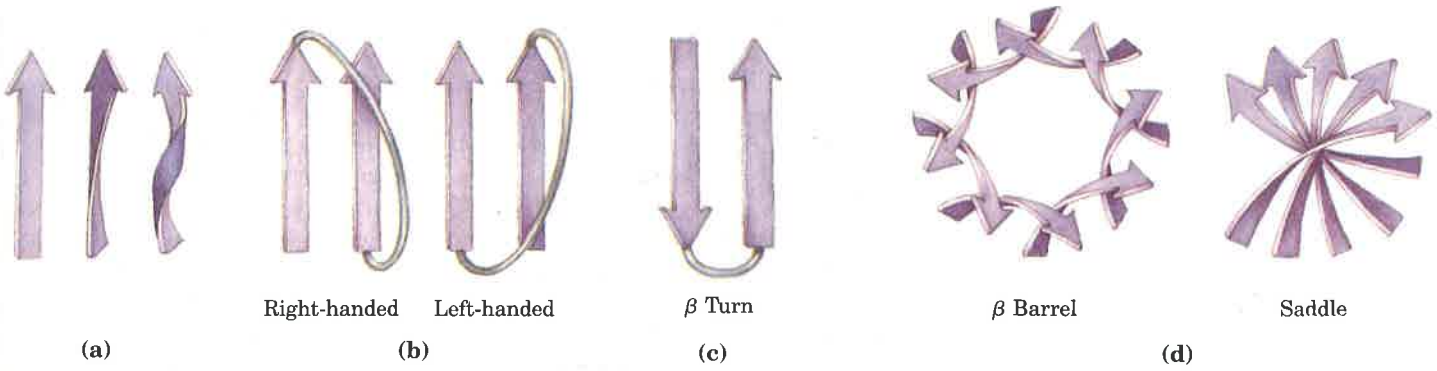



**Figure 7–23** Extended β chains of amino acids tend to twist in a right-handed sense because the slightly twisted conformation is more stable than the linear conformation **(a).** This influences the conformation of the polypeptide segments that connect two β strands, and also the stable conformations assumed by several adjacent β strands. **(b)** Connections between parallel β chains are right-handed. **(c)** The β turn is a common connector between antiparallel β chains. **(d)** The tendency for right-handed twisting is seen in two particularly stable arrangements of adjacent β chains: the β barrel and the saddle; these structures form the stable core of many proteins.

was complete. In an alternative model, folding is initiated by a spontaneous collapse of the polypeptide into a compact state mediated by hydrophobic interactions among nonpolar residues. The state resulting from this "hydrophobic collapse" may have a high content of secondary structure, but many amino acid side chains are not entirely fixed. Either or both models (and perhaps others) may apply to a given protein.

A number of structural constraints help to guide the interaction of regions of secondary structure. The most common patterns are sometimes referred to as supersecondary structures. A prominent one is a tendency for extended β conformations to twist in a right-handed sense (Fig. 7–23a). This influences both the arrangement of β sheets relative to one another and the path of the polypeptide segment connecting two β strands. Two parallel β strands, for example, must be connected by a crossover strand (Fig. 7–23b). In principle, this crossover could have a right- or left-handed conformation, but only the right-handed form is found in proteins. The twisting of β sheets also leads to a characteristic twisting of the structure formed when many sheets are put together. Two examples of resulting structures are the β barrel and saddle shapes (Fig. 7–23d), which form the core of many larger structures.

Weak-bonding interactions represent the ultimate thermodynamic constraint on the interaction of different regions of secondary structure. The R groups of amino acids project outward from α-helical and β structures, and thus the need to bury hydrophobic residues means that water-soluble proteins must have more than one layer of secondary structure. One simple structural method for burying hydrophobic residues is a supersecondary structural unit called a βαβ loop (Fig. 7–24), a structure often repeated multiple times in larger proteins. More elaborate structures are domains made up of facing β sheets (with hydrophobic residues sandwiched between), and β sheets covered on one side with several α helices, as described later.

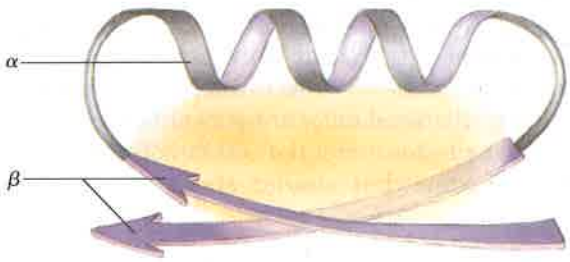



**Figure 7–24** The βαβ loop. The shaded region denotes the area where stabilizing hydrophobic interactions occur.

It becomes more difficult to bury hydrophobic residues in smaller structures, and the number of potential weak interactions available for stabilization decreases. For this reason, smaller proteins are often held together with a number of covalent bonds, principally disulfide linkages. Recall the multiple disulfide bonds in the small proteins insulin (see Fig. 6–10) and ribonuclease (Fig. 7–21). Other types of covalent bonds also occur. The heme group in cytochrome *c*, for example, is covalently linked to the protein on two sides, providing a significant stabilization of the entire protein structure.

Not all proteins fold spontaneously as they are synthesized in the cell. Proteins that facilitate the folding of other proteins have been found in a wide variety of cells. These are called **polypeptide chain binding proteins** or molecular chaperones. Several of these proteins can bind to polypeptide chains, preventing nonspecific aggregation of weak-bonding side chains. They guide the folding of some polypeptides, as well as the assembly of multiple polypeptides into larger structures. Dissociation of polypeptide chain binding proteins from polypeptides is often coupled to ATP hydrolysis. One family of such proteins has structures that are highly conserved in organisms ranging from bacteria to mammals. These proteins ($M_r$ 70,000), as well as several other families of polypeptide chain binding proteins, were originally identified as "heat shock" proteins because they are induced in many cells when heat stress is applied, and apparently help stabilize other proteins.

Some proteins have also been found that promote polypeptide folding by catalyzing processes that otherwise would limit the rate of folding, such as the reversible formation of disulfide bonds or proline isomerization (the interconversion of the cis and trans isomers of peptide bonds involving the imino nitrogen of proline; see Fig. 7–10).

### There Are a Few Common Tertiary Structural Patterns

Following the folding patterns outlined above and others yet to be discovered, a newly synthesized polypeptide chain quickly assumes its most stable tertiary structure. Although each protein has a unique structure, several patterns of tertiary structure seem to occur repeatedly in proteins that differ greatly in biological function and amino acid sequence (Fig. 7–25). This may reflect an unusual degree of stability and/or functional flexibility conferred by these particular tertiary structures. It also demonstrates that biological function is determined not only by the overall three-dimensional shape of the protein, but also by the arrangement of amino acids within that shape.

One structural motif is made up of eight $\beta$ strands arranged in a circle with each $\beta$ strand connected to its neighbor by an $\alpha$ helix. The $\beta$ regions are arranged in the barrel structure described in Figure 7–23, and they influence the overall tertiary structure, giving rise to the name $\alpha/\beta$ barrel (Fig. 7–25a). This structure is found in many enzymes; a binding site for a cofactor or substrate is often found in a pocket formed near an end of the barrel.

Another structural motif is the four-helix bundle (Fig. 7–25b), in which four $\alpha$ helices are connected by three peptide loops. The helices are slightly tilted to form a pocket in the middle, which often contains a binding site for a metal or other cofactors essential for biological function. A somewhat similar structure in which seven helices are ar-

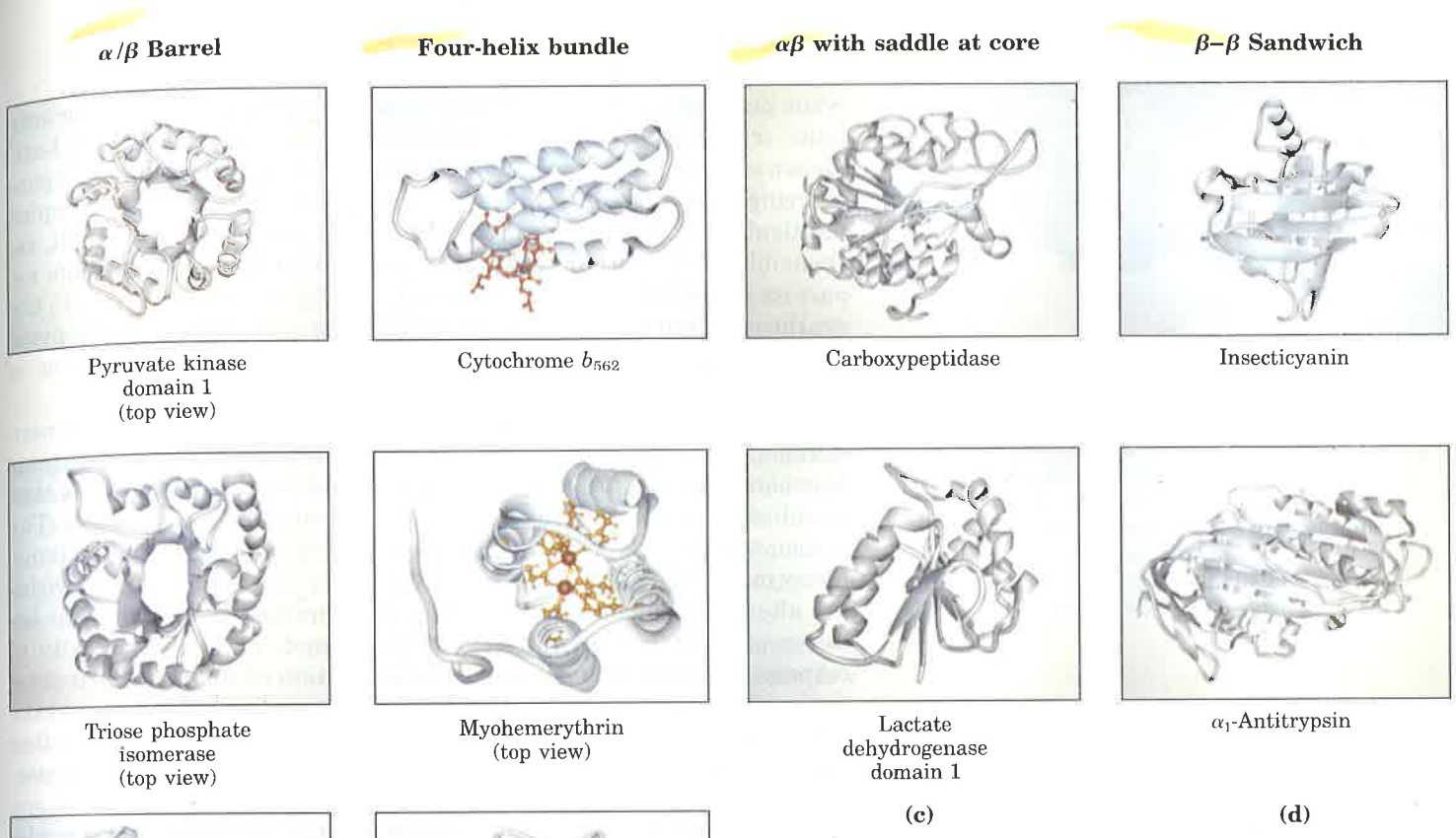



**Figure 7–25** Examples of some common structural motifs in proteins. **(a)** The $\alpha/\beta$ barrel, found in pyruvate kinase and triose phosphate isomerase, enzymes of the glycolytic pathway. This structure also occurs in the larger domain of ribulose-1,5-bisphosphate carboxylase/oxygenase (known also as rubisco), an enzyme essential to the fixation of $CO_2$ by plants; in glycolate oxidase, an enzyme in photorespiration; and in a number of other unrelated proteins. **(b)** The four-helix bundle, shown here in cytochrome $b_{562}$ and myohemerythrin. A dinuclear iron center and coordinating amino acids in myohemerythrin are shown in orange. Myohemerythrin is a nonheme oxygen-transporting protein found in certain worms and mollusks. The four-helix bundle is also found in apoferritin and the tobacco mosaic virus coat-protein. Apoferritin is a widespread protein involved in iron transport and storage. **(c)** $\alpha\beta$ with saddle at core, in carboxypeptidase, a protein-hydrolyzing (proteolytic) enzyme, and lactate dehydrogenase, a glycolytic enzyme. **(d)** $\beta$–$\beta$ Sandwich. In the protein insecticyanin of moths, the hydrophobic pocket binds biliverdin, a colored substance that plays a role in camouflage. $\alpha_1$-Antitrypsin is a naturally occurring inhibitor of the proteolytic enzyme trypsin.

ranged in a barrel-like motif is found in some membrane proteins (see Fig. 10–10). The seven helices often surround a channel that spans the membrane.

A third motif has a $\beta$ sheet in the "saddle" conformation forming a stable core, often surrounded by a number of $\alpha$-helical regions (Fig. 7–25c). Structures of this kind are found in many enzymes. The location of the substrate binding site varies, determined by the placement of the $\alpha$ helices and other variable structural elements.

One final motif makes use of a sandwich of $\beta$ sheets, layered so that the strands of the sheets form a quiltlike cross-hatching when viewed from above (Fig. 7–25d). This creates a hydrophobic pocket between the $\beta$ sheets that is often a binding site for a planar hydrophobic molecule.

## Protein Quaternary Structure

Some proteins contain two or more separate polypeptide chains or subunits, which may be identical or different in structure. One of the best-known examples of a multisubunit protein is hemoglobin, the oxygen-carrying protein of erythrocytes. Among the larger, more complex multisubunit proteins are the enzyme RNA polymerase of *E. coli,* responsible for initiation and synthesis of RNA chains; the enzyme aspartate transcarbamoylase (12 chains; see Fig. 8–26), important in the synthesis of nucleotides; and, as an extreme case, the enormous pyruvate dehydrogenase complex of mitochondria, which is a cluster of three enzymes containing a total of 102 polypeptide chains.

The arrangement of proteins and protein subunits in three-dimensional complexes constitutes quaternary structure. The interactions between subunits are stabilized and guided by the same forces that stabilize tertiary structure: multiple noncovalent interactions. The association of polypeptide chains can serve a variety of functions. Many multisubunit proteins serve regulatory functions; their activities are altered by the binding of certain small molecules. Interactions between subunits can permit very large changes in enzyme activity in response to small changes in the concentration of substrate or regulatory molecules (Chapter 8). In other cases, separate subunits can take on separate but related functions. Entire metabolic pathways are often organized by the association of a supramolecular complex of enzymes, permitting an efficient channeling of pathway intermediates from one enzyme to the next. Other associations, such as the histones in a nucleosome or the coat proteins of a virus, serve primarily structural roles. Large assemblies sometimes reflect complex functions. One obvious example is the complicated structure of ribosomes (see Fig. 26–12), which carry out protein synthesis.

X-ray and other analytical methods for structure determination become more difficult as the size and number of subunits in a protein increases. Nevertheless, sufficient data are already available to yield some very important information about the structure and function of multisubunit proteins.

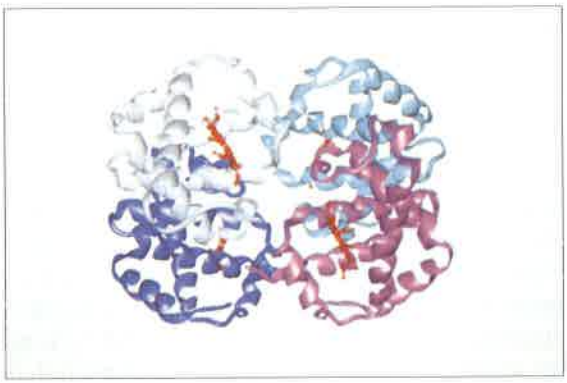
Max Perutz

John Kendrew

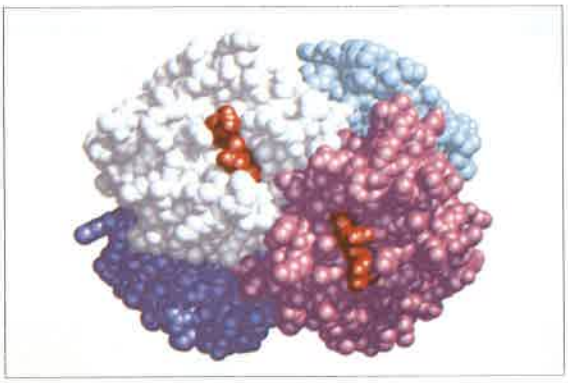
### X-Ray Analysis Revealed the Complete Structure of Hemoglobin

The first oligomeric protein to be subjected to x-ray analysis was hemoglobin ($M_r$ 64,500), which contains four polypeptide chains and four heme prosthetic groups, in which the iron atoms are in the ferrous ($Fe^{2+}$) state. The protein portion, called globin, consists of two $\alpha$ chains (141 residues each) and two $\beta$ chains (146 residues each). Note that $\alpha$ and $\beta$ do not refer to secondary structures in this case. Because hemoglobin is four times as large as myoglobin, much more time and effort were required to solve its three-dimensional structure, finally achieved by Max Perutz, John Kendrew, and their colleagues in 1959.

The hemoglobin molecule is roughly spherical, with a diameter of about 5.5 nm. The $\alpha$ and $\beta$ chains contain several segments of $\alpha$ helix separated by bends, with a tertiary structure very similar to that of the single polypeptide of myoglobin. In fact, there are 27 invariant amino acid residues in these three polypeptide chains, and closely related amino acids at 40 additional positions, indicating that these polypeptides (myoglobin and the $\alpha$ and $\beta$ chains of hemoglobin) are evolutionarily related. The four polypeptide chains in hemoglobin fit together in an approximately tetrahederal arrangement (Fig. 7–26).



(a)



(b)

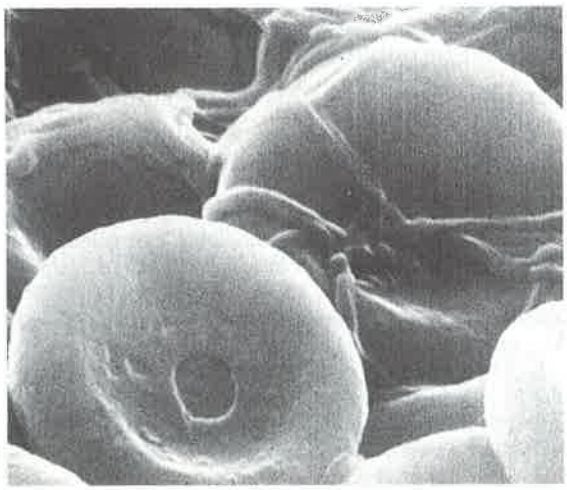
**Figure 7–26** The three-dimensional (quaternary) structure of deoxyhemoglobin, revealed by x-ray diffraction analysis, showing how the four subunits are packed together. **(a)** A ribbon representation. **(b)** A space-filling model. The $\alpha$ subunits are shown in white and light blue; the $\beta$ subunits are shown in pink and purple. Note that the heme groups, shown in red, are relatively far apart.

One heme is bound to each polypeptide chain of hemoglobin. The oxygen-binding sites are rather far apart given the size of the molecule, about 2.5 nm from one another. Each heme is partially buried in a pocket lined with hydrophobic amino acid side chains. It is bound to its polypeptide chain through a coordination bond of the iron atom to the R group of a His residue (see Fig. 7–18). The sixth coordination bond of the iron atom of each heme is available to bind $O_2$.

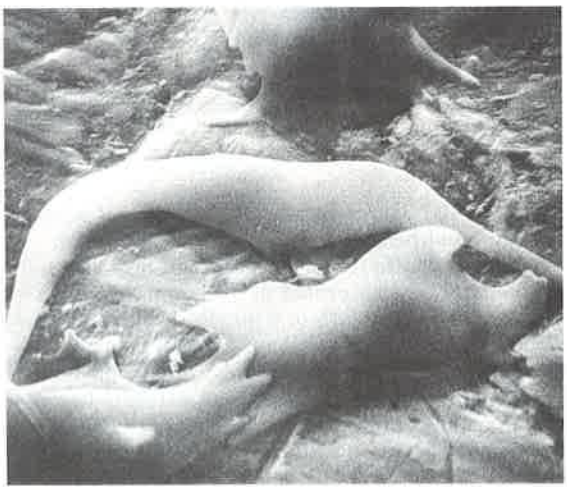
Closer examination of the quaternary structure of hemoglobin, with the help of molecular models, shows that although there are few contacts between the two $\alpha$ chains or between the two $\beta$ chains, there are many contact points between the $\alpha$ and $\beta$ chains. These contact points consist largely of hydrophobic side chains of amino acid residues, but also include ionic interactions involving the carboxyl-terminal residues of the four subunits.

Naturally occurring changes in the amino acid sequence of hemoglobin provide some useful insights into the relationship between structure and function in proteins. More than 300 genetic variants of hemoglobin are known to occur in the human population. Most of these variations are single amino acid changes that have only minor structural or functional effects. An exception is a substitution of valine for glutamate at position 6 of the $\beta$ chain. This residue is on the outer surface of the molecule, and the change produces a "sticky" hydrophobic spot on the surface that results in abnormal quaternary association of hemoglobin. When oxygen concentrations are below a critical level, the subunits polymerize into linear arrays of fibers that distort cell shape. The result is a sickling of erythrocytes (Fig. 7–27), the cause of sickle-cell anemia.

**Figure 7–27** Scanning electron micrographs of **(a)** normal and **(b)** sickled human erythrocytes. The sickled cells are fragile, and their breakdown causes anemia.



(a)



(b)

## Conformational Changes in Hemoglobin Alter Its Oxygen-Binding Capacity

Hemoglobin is an instructive model for studying the function of many regulatory oligomeric proteins. The blood in a human being must carry about 600 L of oxygen from the lungs to the tissues every day, but very little of this is carried by the blood plasma because oxygen is only sparingly soluble in aqueous solutions. Nearly all the oxygen carried by whole blood is bound and transported by the hemoglobin of the erythrocytes. Normal human erythrocytes are small (6 to 9 $\mu$m), biconcave disks (Fig. 7–27a). They have no nucleus, mitochondria, endoplasmic reticulum, or other organelles. The hemoglobin of the erythrocytes in arterial blood passing from the lungs to the peripheral tissues is about 96% saturated with oxygen. In the venous blood returning to the heart, the hemoglobin is only about 64% saturated. Thus blood passing through a tissue releases about one-third of the oxygen it carries.

The special properties of the hemoglobin molecule that make it such an effective oxygen carrier are best understood by comparing the $O_2$-binding or $O_2$-saturation curves of myoglobin and hemoglobin (Fig. 7–28). These show the percentage of $O_2$-binding sites of hemoglobin or myoglobin that are occupied by $O_2$ molecules when solutions of these proteins are in equilibrium with different partial pressures of oxygen in the gas phase. (The partial pressure of oxygen, abbreviated $pO_2$, is the pressure contributed by oxygen to the overall pressure of a mixture of gases, and is directly related to the concentration of oxygen in the mixture.)

From its saturation curve, it is clear that myoglobin has a very high affinity for oxygen (Fig. 7–28). Furthermore, the $O_2$-saturation curve of myoglobin is a simple hyperbolic curve, as might be expected from the mass action of oxygen on the equilibrium myoglobin + $O_2 \rightleftharpoons$ oxymyoglobin. In contrast, the oxygen affinity of each of the four $O_2$-binding sites of deoxyhemoglobin is much lower, and the $O_2$-saturation curve of hemoglobin is sigmoid (S-shaped) (Fig. 7–28). This shape indicates that whereas the affinity of hemoglobin for binding the first $O_2$ molecule (to any of the four sites) is relatively low, the second, third, and fourth $O_2$ molecules are bound with a very much higher affinity. This accounts for the steeply rising portion of the sigmoid curve. The increase in the affinity of hemoglobin for oxygen after the first $O_2$ molecule is bound is almost 500-fold. Thus the oxygen affinity of each heme–polypeptide subunit of hemoglobin depends on whether $O_2$ is bound to neighboring subunits. The conversion of deoxyhemoglobin to oxyhemoglobin requires the disruption of ionic interactions involving the carboxyl-terminal residues of the four subunits, interactions that con-

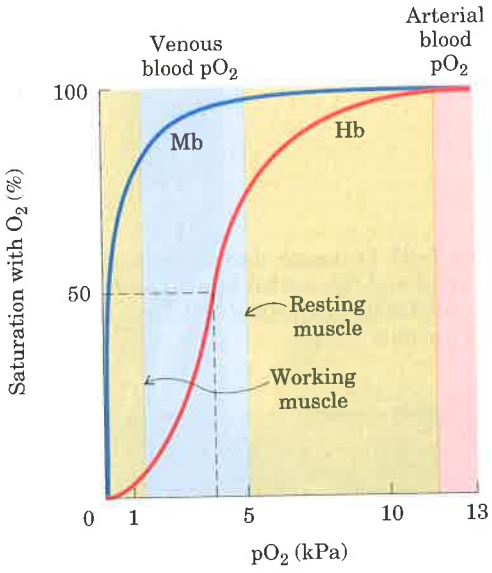



**Figure 7–28** The oxygen-binding curves of myoglobin (Mb) and hemoglobin (Hb). Myoglobin has a much greater affinity for oxygen than does hemoglobin. It is 50% saturated at oxygen partial pressures ($pO_2$) of only 0.15 to 0.30 kPa, whereas hemoglobin requires a $pO_2$ of about 3.5 kPa for 50% saturation. Note that although both hemoglobin and myoglobin are more than 95% saturated at the $pO_2$ in arterial blood leaving the lungs (~13 kPa), hemoglobin is only about 75% saturated in resting muscle, where the $pO_2$ is about 5 kPa, and only 10% saturated in working muscle, where the $pO_2$ is only about 1.5 kPa. Thus hemoglobin can release its oxygen very effectively in muscle and other peripheral tissues. Myoglobin, on the other hand, is still about 80% saturated at a $pO_2$ of 1.5 kPa, and therefore unloads very little oxygen even at very low $pO_2$. Thus the sigmoid $O_2$-saturation curve of hemoglobin is a molecular adaptation for its transport function in erythrocytes, assuring the binding and release of oxygen in the appropriate tissues.

strain the overall structure in a low-affinity state. The increase in affinity for successive $O_2$ molecules reflects the fact that more of these ionic interactions must be broken for binding the first $O_2$ than for binding later ones.

Once the first heme–polypeptide subunit binds an $O_2$ molecule, it communicates this information to the remaining subunits through interactions at the subunit interfaces. The subunits respond by greatly increasing their oxygen affinity. This involves a change in the conformation of hemoglobin that occurs when oxygen binds (Fig. 7–29). Such communication among the four heme–polypeptide subunits of hemoglobin is the result of cooperative interactions among the subunits. Because binding of one $O_2$ molecule increases the probability that further $O_2$ molecules will be bound by the remaining subunits, hemoglobin is said to have **positive cooperativity.** Sigmoid binding curves, like that of hemoglobin for oxygen, are characteristic of positive cooperative binding. Cooperative oxygen binding does not occur with myoglobin, which has only one heme group within a single polypeptide chain and thus can bind only one $O_2$ molecule; its saturation curve is therefore hyperbolic. The multiple subunits of hemoglobin and the interactions between these subunits result in a fundamental difference between the $O_2$-binding actions of myoglobin and hemoglobin.

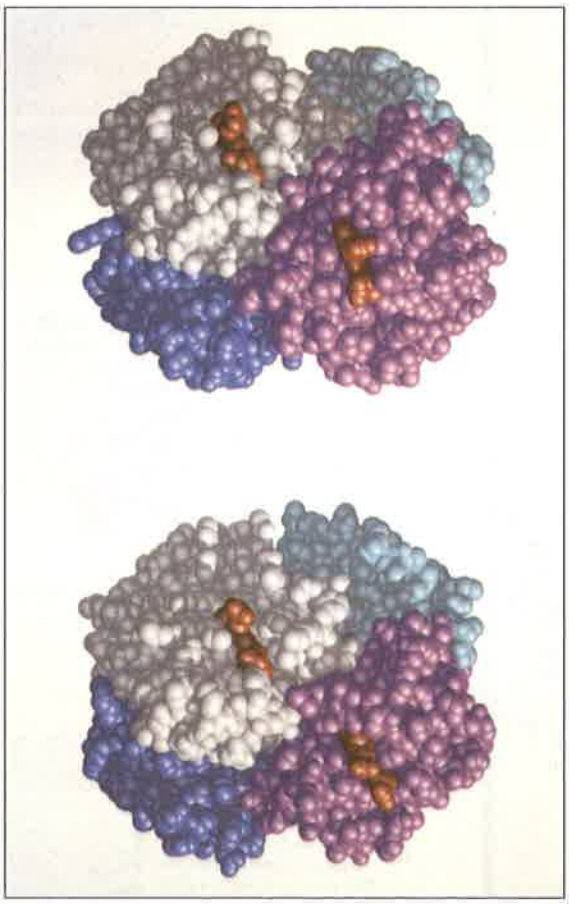
Positive cooperativity is not the only result of subunit interactions in oligomeric proteins. Some oligomeric proteins show **negative cooperativity:** binding of one ligand molecule *decreases* the probability that further ligand molecules will be bound. These and additional regulatory mechanisms used by these proteins are considered in Chapter 8.

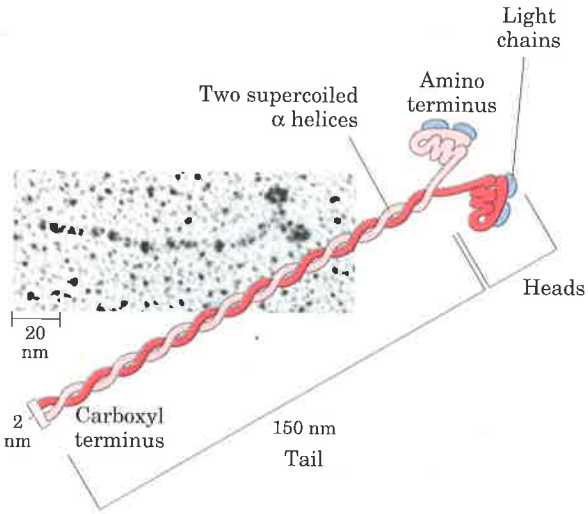
## Hemoglobin Binds Oxygen in the Lungs and Releases It in Peripheral Tissues

In the lungs the $pO_2$ in the air spaces is about 13 kPa; at this pressure hemoglobin is about 96% saturated with oxygen. However, in the cells of a working muscle the $pO_2$ is only about 1.5 kPa because muscle cells use oxygen at a high rate and thus lower its local concentration. As the blood passes through the muscle capillaries, oxygen is released from the nearly saturated hemoglobin in the erythrocytes into the blood plasma and thence into the muscle cells. As is evident from the $O_2$-saturation curve in Figure 7–28, hemoglobin releases about a third of its bound oxygen as it passes through the muscle capillaries, so that when it leaves the muscle, it is only about 64% saturated. When the blood returns to the lungs, where the $pO_2$ is much higher (13 kPa), the hemoglobin quickly binds more oxygen until it is 96% saturated again.

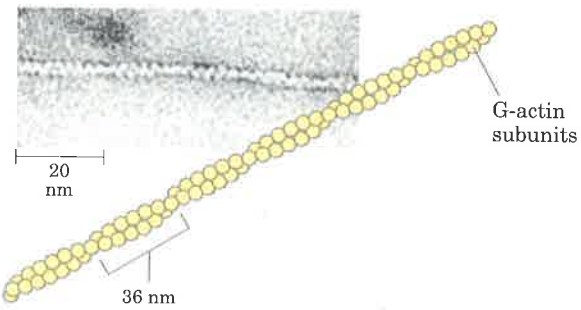
Now suppose that the hemoglobin in the erythrocyte were replaced by myoglobin. We see from the hyperbolic $O_2$-saturation curve of myoglobin (Fig. 7–28) that only 1 or 2% of the bound oxygen can be released from myoglobin as the $pO_2$ decreases from 13 kPa in the lungs to 3 kPa in the muscle. Myoglobin therefore is not very well adapted for carrying oxygen from the lungs to the tissues, because it has a much higher affinity for oxygen and releases very little of it at the $pO_2$ in muscles and other peripheral tissues. However, in its true biological function *within* muscle cells, which is to store oxygen and make it available to the mitochondria, myoglobin is in fact much better suited than hemoglobin, because its very high affinity for oxygen at low $pO_2$ enables it to bind and store oxygen effectively. Thus hemoglobin and myoglobin are specialized and adapted for different kinds of $O_2$-binding functions.

**Figure 7–29** Conformational changes induced in hemoglobin when oxygen binds. (The oxygen-bound form is shown at bottom.) There are multiple structural changes, some not visible here; most of the changes are subtle. The $\alpha$ and $\beta$ subunits are colored as in Fig. 7–26.






(a)




(b)

**Figure 7–30** Myosin and actin, the two filamentous proteins of contractile systems. **(a)** The myosin molecule has a long tail consisting of two supercoiled α-helical polypeptide chains (heavy chains). The head of each heavy chain is associated with two light chains and is an enzyme capable of hydrolyzing ATP. **(b)** A representation of an F-actin fiber, which consists of two chains of G-actin subunits coiled about each other to form a filament.

## There Are Limits to the Size of Proteins

The relatively large size of proteins reflects their functions. The function of an enzyme, for example, requires a protein large enough to form a specifically structured pocket to bind its substrate. The size of proteins has limits, however, imposed by the genetic coding capacity of nucleic acids and the accuracy of the protein biosynthetic process. The use of many copies of one or a few proteins to make a large enclosing structure is important for viruses because this strategy conserves genetic material. Remember that there is a linear correspondence between the sequence of a gene in nucleic acid and the amino acid sequence of the protein for which it codes (see Fig. 6–14). The nucleic acids of viruses are much too small to encode the information required for a protein shell made of a single polypeptide. By using many copies of much smaller proteins for the virus coat, a much shorter nucleic acid is needed for the protein subunits, and this nucleic acid can be efficiently used over and over again. Cells also use large protein complexes in muscle, cilia, the cytoskeleton, and other structures. It is simply more efficient to make many copies of a small protein than one copy of a very large one. The second factor limiting the size of proteins is the error frequency during protein biosynthesis. This error frequency is low but can become significant for very large proteins. Simply put, the potential for incorporating a "wrong" amino acid in a protein is greater for a large protein than a small one.

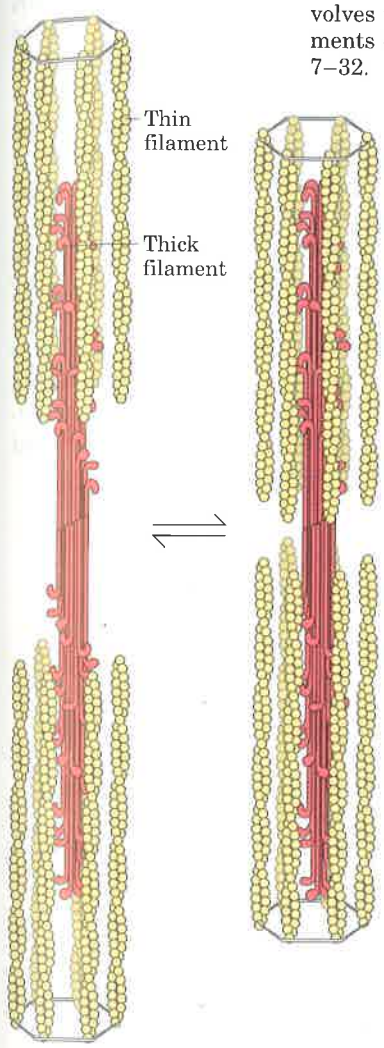
## Some Proteins Form Supramolecular Complexes

The same principles that govern the stability of secondary, tertiary, and quaternary structure in proteins guide the formation of very large protein complexes. These function, for example, as biological engines (muscle and cilia), large structural enclosures (virus coats), cellular skeletons (actin and tubulin filaments), DNA-packaging complexes (chromatin), and machines for protein synthesis (ribosomes). In many cases the complex consists of a small number of distinct proteins, specialized so that they spontaneously polymerize to form large structures.

Muscle provides an example of a supramolecular complex of multiple copies of a limited number of proteins. The contractile force of muscle is generated by the interaction of two proteins, actin and myosin (Chapter 2). Myosin is a long, rodlike molecule ($M_r$ 540,000) consisting of six polypeptide chains, two so-called heavy chains ($M_r$ ~230,000) and four light chains ($M_r$ ~20,000) (Fig. 7–30a). The two heavy chains have long α-helical tails that twist around each other in a left-handed fashion. The large head domain, at one end of each heavy chain, interacts with actin and contains a catalytic site for ATP hydrolysis. Many myosin molecules assemble together to form the **thick filaments** of skeletal muscle (Fig. 7–31).

The other protein, actin, is a polymer of the globular protein G-actin ($M_r$ 42,000); two such polymers coil around each other in a right-handed helix to form a **thin filament** (Fig. 7–30b). The interaction between actin and myosin is dynamic; contacts consist of multiple weak interactions that are strong enough to provide a stable association but weak enough to allow dissociation when needed. Hydrolysis of ATP in the myosin head is coupled to a series of conformational changes that bring about muscle contraction (Fig. 7–32). A similar engine involving an interaction between tubulin and dynein brings about the motion of cilia (Chapter 2).

**Figure 7–31** The thick and thin filaments of muscle. Many myosin molecules assemble in a bundle to form a thick filament. Muscle contraction involves the sliding of thick filaments past thin filaments of actin, by a mechanism described in Fig. 7–32.



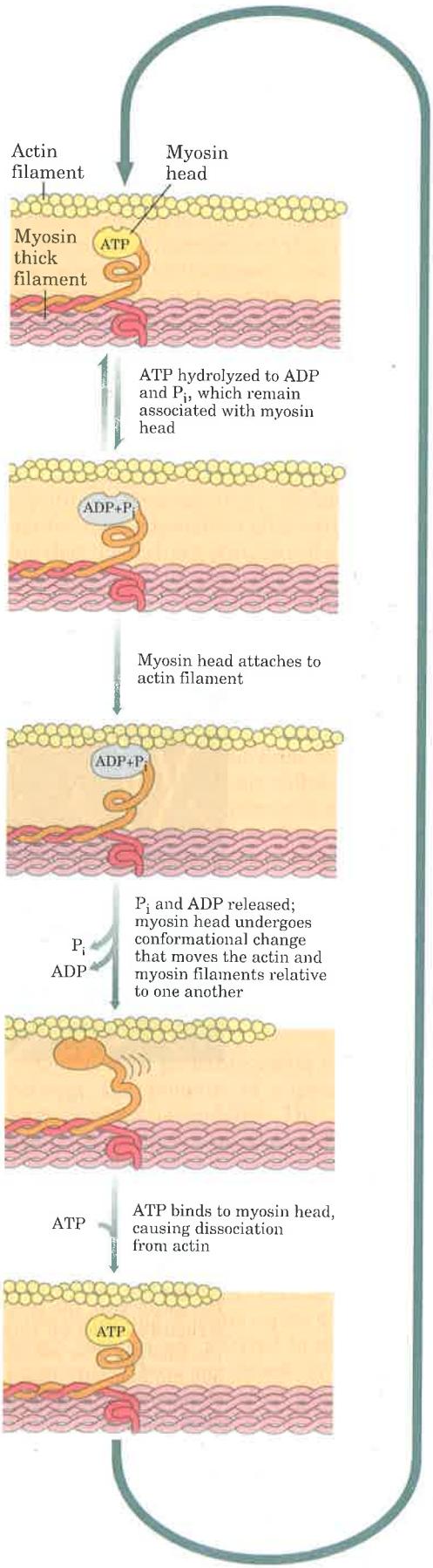




**Figure 7–32** The sliding of the thick and thin filaments in muscle involves an interaction between actin and myosin mediated by ATP hydrolysis ($ATP \rightleftharpoons ADP + P_i$, where $P_i$ is inorganic phosphate, or $PO_4^{3-}$). Conformational changes in the myosin head that are coupled to stages in the ATP hydrolytic cycle cause myosin to successively dissociate from one actin subunit and then associate with another farther along the actin filament. In this way the myosin heads "walk" along the thin filaments, and draw the thin filament array into the thick filament array.

The protein structures in virus coats (called capsids) generally function simply as enclosures. In many cases capsids are made up of one or a few proteins that assemble spontaneously around a viral DNA or RNA molecule. Two types of viral structures are shown in Figure 7–33. The tobacco mosaic virus is a right-handed helical filament with 2,130 copies of a single protein that interact to form a cylinder enclosing the RNA genome. Another common structure for virus coats is the icosahedron, a regular 12-cornered polyhedron having 20 equilateral triangular faces. Two examples are poliovirus and human rhinovirus 14 (a common cold virus), each made up of 60 protein units (Fig. 7–33). Each protein unit consists of single copies of four different polypeptide chains, three of which are accessible at the outer surface. The resulting shell encloses the genetic material (RNA) of the virus.

The primary forces guiding the assembly of even these very large structures are the weak noncovalent interactions that have dominated this discussion. Each protein has several surfaces that are complementary to surfaces in adjacent protein subunits. Each protein is most stable only when it is part of the larger structure.

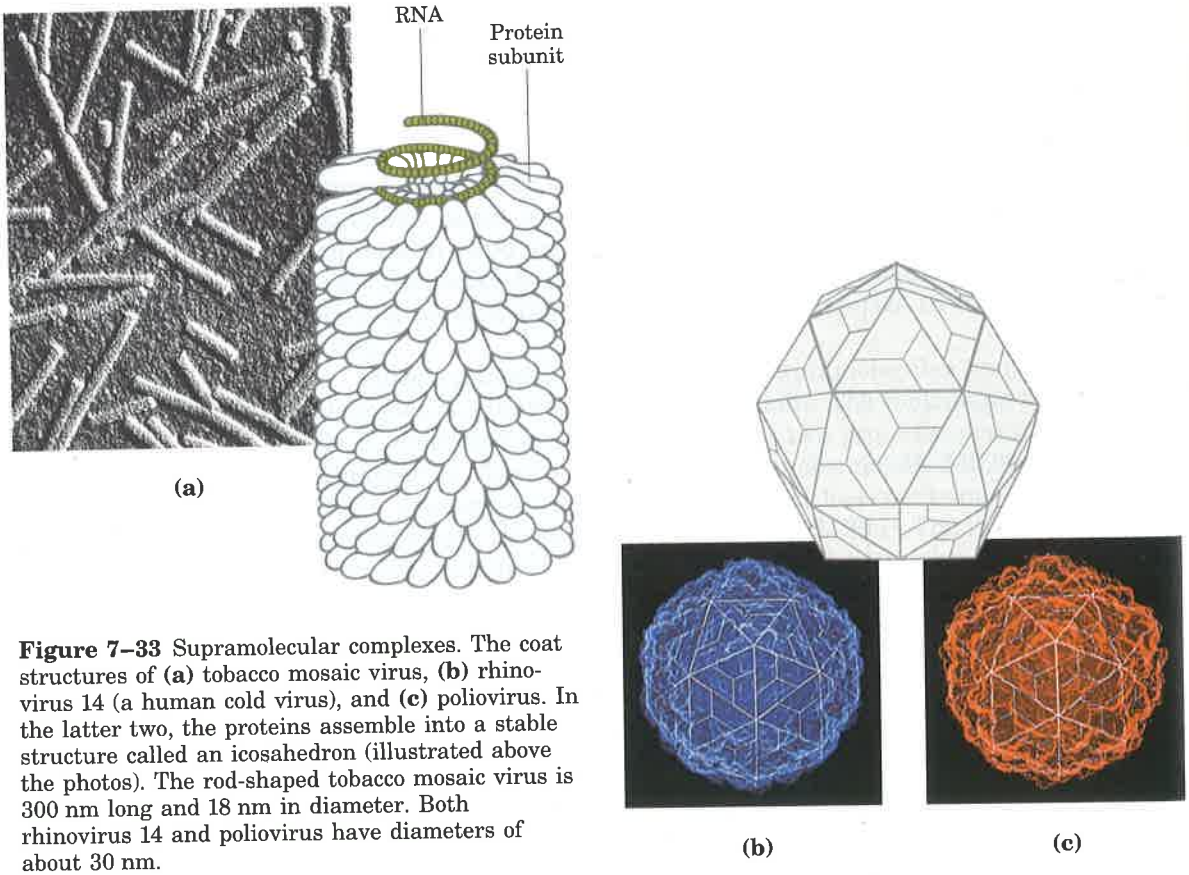



**Figure 7–33** Supramolecular complexes. The coat structures of **(a)** tobacco mosaic virus, **(b)** rhinovirus 14 (a human cold virus), and **(c)** poliovirus. In the latter two, the proteins assemble into a stable structure called an icosahedron (illustrated above the photos). The rod-shaped tobacco mosaic virus is 300 nm long and 18 nm in diameter. Both rhinovirus 14 and poliovirus have diameters of about 30 nm.

## Summary

Every protein has a unique three-dimensional structure that reflects its function, a structure stabilized by multiple weak interactions. Hydrophobic interactions provide the major contribution to stabilizing the globular form of most soluble proteins; hydrogen bonds and ionic interactions are optimized in the specific structure that is thermodynamically most stable.

There are four generally recognized levels of protein structure. Primary structure refers to the amino acid sequence and the location of disulfide bonds. Secondary structure refers to the spatial relationship of adjacent amino acids. Tertiary structure is the three-dimensional conformation of an entire polypeptide chain. Quaternary structure involves the spatial relationship of multiple polypeptide chains (e.g., enzyme subunits) that are tightly associated.

The nature of the bonds in the polypeptide chain places constraints on structure. The peptide bond is characterized by a partial double-bond character that keeps the entire amide group in a rigid planar configuration. The N—$C_\alpha$ and $C_\alpha$—C bonds can rotate with bond angles $\phi$ and $\psi$, respectively. Secondary structure can be defined completely by these two bond angles.

There are two general classes of proteins: fibrous and globular. Fibrous proteins, which serve mainly structural roles, have simple repeating structures and provided excellent models for the early studies of protein structure. Two major types of secondary structure were predicted by model building based on information obtained from fibrous proteins: the $\alpha$ helix and the $\beta$ conformation. Both are characterized by optimal hydrogen bonding between amide nitrogens and carbonyl oxygens in the peptide backbone. The stability of these structures within a protein is influenced by their amino acid content and by the relative placement of amino acids in the sequence. Another nonrepeating type of secondary structure common in proteins is the $\beta$ bend.

In fibrous proteins such as keratin and collagen, a single type of secondary structure predominates. The polypeptide chains are supertwisted into ropes and then combined in larger bundles to provide strength. The structure of elastin permits stretching.

Globular proteins have more complicated tertiary structures, often containing several types of secondary structure in the same polypeptide chain. The first globular protein structure to be determined, using x-ray diffraction methods, was that of myoglobin. This structure confirmed that a predicted secondary structure ($\alpha$ helix) occurs in proteins; that hydrophobic amino acids are located in the protein interior; and that globular proteins are compact. Subsequent research on protein structure has reinforced these conclusions while demonstrating that different proteins often differ in tertiary structure.

The three-dimensional structure of proteins can be destroyed by treatments that disrupt weak interactions, a process called denaturation. Denaturation destroys protein function, demonstrating a relationship between structure and function. Some denatured proteins (e.g., ribonuclease) can renature spontaneously to give active protein, showing that the tertiary structure of a protein is determined by its amino acid sequence.

The folding of globular proteins is believed to begin with local formation of regions of secondary structure, followed by interactions of these regions and adjustments to reach the final tertiary structure. Sometimes regions of a polypeptide chain, called domains, fold up separately and can have separate functions. The final structure and the steps taken to reach it are influenced by the need to bury hydrophobic amino acid side chains in the protein interior away from water, the tendency of a polypeptide chain to twist in a right-handed sense, and the need to maximize hydrogen bonds and ionic interactions. These constraints give rise to structural patterns such as the $\beta\alpha\beta$ fold and twisted $\beta$ pleated sheets. Even at the level of tertiary structure, some common patterns are found in proteins that have no known functional relationship.

Quaternary structure refers to the interaction between the subunits of oligomeric proteins or large protein assemblies. The best-studied oligomeric protein is hemoglobin. The four subunits of hemoglobin exhibit cooperative interactions on oxygen binding. Binding of oxygen to one subunit facilitates oxygen binding to the next, giving rise to a sigmoid binding curve. These effects are mediated by subunit–subunit interactions and subunit conformational changes. Very large protein structures consisting of many copies of one or a few different proteins are referred to as supramolecular complexes. These are found in cellular skeletal structures, muscle and other types of cellular "engines," and virus coats.

## Further Reading

### *General*

**Anfinsen, C.B.** (1973) Principles that govern the folding of protein chains. *Science* **181,** 223–230.

*The author reviews his classic work on ribonuclease.*

**Cantor, C.R. & Schimmel, P.R.** (1980) *Biophysical Chemistry,* Part I: *The Conformation of Biological Macromolecules,* W.H. Freeman and Company, New York.

Evolution of Catalytic Function. (1987) *Cold Spring Harb. Symp. Quant. Biol.* **52.**

*A source of excellent articles on many topics, including protein structure, folding, and function.*

**Creighton, T.E.** (1984) *Proteins: Structures and Molecular Properties,* W.H. Freeman and Company, New York.

**Oxender, D.L. (ed)** (1987) *Protein Structure, Folding, and Design 2,* UCLA Symposia on Molecular and Cellular Biology, New Series, Vol. 69, Alan R. Liss, Inc., New York.

*Summary papers from a major symposium on the title subject.*

Structure and Function of Proteins. (1989) *Trends Biochem. Sci.* **14** (July).

*A special issue devoted to reviews on protein chemistry and protein structure. Includes good summaries of protein folding, protein structure prediction, and many other topics.*

### *Secondary, Tertiary, and Quaternary Structure*

**Dickerson, R.E. & Geis, I.** (1982) *Hemoglobin: Structure, Function, Evolution, and Pathology,* The Benjamin/Cummings Publishing Company, Menlo Park, CA.

**Ingram, V.M.** (1957) Gene mutations in human haemoglobin: the chemical difference between normal and sickle cell haemoglobin. *Nature* **180,** 326–328.

*Discovery of the amino acid replacement in sickle-cell hemoglobin (hemoglobin S).*

**Kendrew, J.C.** (1961) The three-dimensional structure of a protein molecule. *Sci. Am.* **205** (December), 96–111.

*Describes how the structure of myoglobin was determined and what was learned from it.*

**Kim, P.S. & Baldwin, R.L.** (1990) Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* **59,** 631–660.

**Koshland, D.E., Jr.** (1973) Protein shape and biological control. *Sci. Am.* **229** (October), 52–64.

*A discussion of the importance of flexibility in protein structures.*

**McPherson, A.** (1989) Macromolecular crystals, *Sci. Am.,* **260** (March), 62–69.

*Describes how macromolecules such as proteins are crystallized.*

**Pace, C.N.** (1990) Conformational stability of globular proteins. *Trends Biochem. Sci.* **15,** 14–17.

**Perutz, M.F.** (1978) Hemoglobin structure and respiratory transport. *Sci. Am.* **239** (December), 92–125.

**Richards, F.M.** (1991) The protein folding problem. *Sci. Am.* **264** (January), 54–63.

**Richardson, J.S.** (1981) The anatomy and taxonomy of protein structure. *Adv. Prot. Chem.* **34,** 167–339.

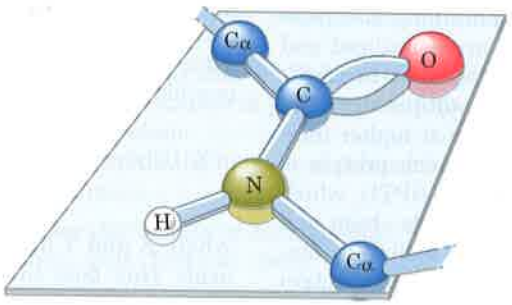
*An outstanding summary of protein structural patterns and principles; the author originated the very useful "ribbon" representations of protein structure that are used in many places in this chapter.*

**Rothman, J.E.** (1989) Polypeptide chain binding proteins: catalysts of protein folding and related processes in cells. *Cell* **59,** 591–601.

**Shortle, D.** (1989) Probing the determinants of protein folding and stability with amino acid substitutions. *J. Biol. Chem.* **264,** 5315–5318.

## Problems

**1.** *Properties of the Peptide Bond* In x-ray studies of crystalline peptides Linus Pauling and Robert Corey found that the C—N bond in the peptide link is intermediate in length (0.132 nm) between a typical C—N single bond (0.149 nm) and a C=N double bond (0.127 nm). They also found that the peptide bond is planar (all four atoms attached to the C—N group are located in the same plane) and that the two $\alpha$-carbon atoms attached to the C—N are always trans to each other (on opposite sides of the peptide bond):




(a) What does the length of the C—N bond in the peptide linkage indicate about its strength and its bond order, i.e., whether it is single, double, or triple?

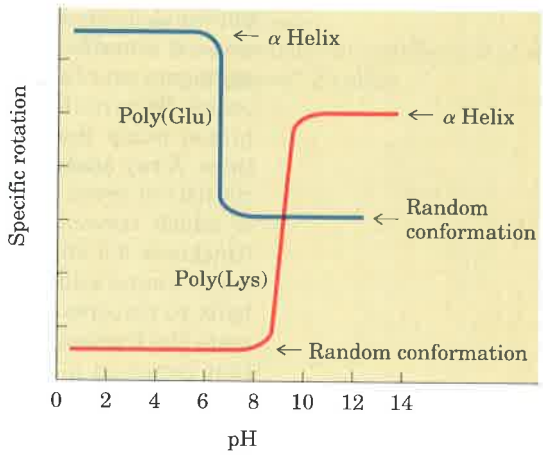
(b) In light of your answer to part (a), provide an explanation for the observation that such a C—N bond is intermediate in length between a double and single bond.

(c) What do the observations of Pauling and Corey tell us about the ease of rotation about the C—N peptide bond?

**2.** *Early Observations on the Structure of Wool* William Astbury discovered that the x-ray pattern of wool shows a repeating structural unit spaced about 0.54 nm along the direction of the wool fiber. When he steamed and stretched the wool, the x-ray pattern showed a new repeating structural unit at a spacing of 0.70 nm. Steaming and stretching the wool and then letting it shrink gave an x-ray pattern consistent with the original spacing of about 0.54 nm. Although these observations provided important clues to the molecular structure of wool, Astbury was unable to interpret them at the time. Given our current understanding of the structure of wool, interpret Astbury's observations.

**3.** *Rate of Synthesis of Hair $\alpha$-Keratin* In human dimensions, the growth of hair is a relatively slow process, occurring at a rate of 15 to 20 cm/yr. All this growth is concentrated at the base of the hair fiber, where $\alpha$-keratin filaments are synthesized inside living epidermal cells and assembled into ropelike structures (see Fig. 7–13). The fundamental structural element of $\alpha$-keratin is the $\alpha$ helix, which has 3.6 amino acid residues per turn and a rise of 0.56 nm per turn (see Fig. 7–6). Assuming that the biosynthesis of $\alpha$-helical keratin chains is the rate-limiting factor in the growth of hair, calculate the rate at which peptide bonds of $\alpha$-keratin chains must be synthesized (peptide bonds per second) to account for the observed yearly growth of hair.

**4.** *The Effect of pH on the Conformations of Polyglutamate and Polylysine* The unfolding of the $\alpha$ helix of a polypeptide to a randomly coiled conformation is accompanied by a large decrease in a property called its specific rotation, a measure of a solution's capacity to rotate plane-polarized light. Polyglutamate, a polypeptide made up of only L-Glu residues, has the $\alpha$-helical conformation at pH 3. However, when the pH is raised to 7, there is a large decrease in the specific rotation of the solution. Similarly, polylysine (L-Lys residues) is an $\alpha$ helix at pH 10, but when the pH is lowered to 7, the specific rotation also decreases, as shown by the following graph.




What is the explanation for the effect of the pH changes on the conformations of poly(Glu) and poly(Lys)? Why does the transition occur over such a narrow range of pH?

**5.** *The Disulfide-Bond Content Determines the Mechanical Properties of Many Proteins* A number of natural proteins are very rich in disulfide

bonds, and their mechanical properties (tensile strength, viscosity, hardness, etc.) are correlated with the degree of disulfide bonding. For example, glutenin, a wheat protein rich in disulfide bonds, is responsible for the cohesive and elastic character of dough made from wheat flour. Similarly, the hard, tough nature of tortoise shell is due to the extensive disulfide bonding in its $\alpha$-keratin. What is the molecular basis for the correlation between disulfide-bond content and mechanical properties of the protein?

**6.** *Why Does Wool Shrink?* When wool sweaters or socks are washed in hot water and/or dried in an electric dryer, they shrink. From what you know of $\alpha$-keratin structure, how can you account for this? Silk, on the other hand, does not shrink under the same conditions. Explain.

**7.** *Heat Stability of Proteins Containing Disulfide Bonds* Most globular proteins are denatured and lose their activity when briefly heated to 65 °C. Globular proteins that contain multiple disulfide bonds often must be heated longer at higher temperatures to denature them. One such protein is bovine pancreatic trypsin inhibitor (BPTI), which has 58 amino acid residues in a single chain and contains three disulfide bonds. On cooling a solution of denatured BPTI, the activity of the protein is restored. Can you suggest a molecular basis for this property?

**8.** *Bacteriorhodopsin in Purple Membrane Proteins* Under the proper environmental conditions, the salt-loving bacterium *Halobacterium halobium* synthesizes a membrane protein ($M_r$ 26,000) known as bacteriorhodopsin, which is purple because it contains retinal. Molecules of this protein aggregate into "purple patches" in the cell membrane. Bacteriorhodopsin acts as a light-activated proton pump that provides energy for cell functions. X-ray analysis of this protein reveals that it consists of seven parallel $\alpha$-helical segments, each of which traverses the bacterial cell membrane (thickness 4.5 nm). Calculate the minimum number of amino acids necessary for one segment of $\alpha$ helix to traverse the membrane completely. Estimate the fraction of the bacteriorhodopsin protein that occurs in $\alpha$-helical form. Justify all your assumptions. (Use an average amino acid residue weight of 110.)

**9.** *Biosynthesis of Collagen* Collagen, the most abundant protein in mammals, has an unusual amino acid composition. Unlike most other proteins, collagen is very rich in proline and hydroxyproline (see p. 172). Hydroxyproline is not one of the 20 standard amino acids, and its incorporation in collagen could occur by one of two routes: (1) proline is hydroxylated by enzymes *before* incorporation into collagen or (2) a Pro residue is hydroxylated *after* incorporation into collagen. To differentiate between these two possibilities, the following experiments were performed. When [$^{14}$C]proline was administered to a rat and the collagen from the tail isolated, the newly synthesized tail collagen was found to be radioactive. If, however, [$^{14}$C]hydroxyproline was administered to a rat, no radioactivity was observed in the newly synthesized collagen. How do these experiments differentiate between the two possible mechanisms for introducing hydroxyproline into collagen?

**10.** *Pathogenic Action of Bacteria That Cause Gas Gangrene* The highly pathogenic anaerobic bacterium *Clostridium perfringens* is responsible for gas gangrene, a condition in which animal tissue structure is destroyed. This bacterium secretes an enzyme that efficiently catalyzes the hydrolysis of the peptide bond indicated in red in the sequence:

$$\text{—X Gly—Pro—Y—} \xrightarrow{H_2O} \text{—X—COO}^- + H_3\overset{+}{N}\text{—Gly—Pro—Y—}$$

where X and Y are any of the 20 standard amino acids. How does the secretion of this enzyme contribute to the invasiveness of this bacterium in human tissues? Why does this enzyme not affect the bacterium itself?

**11.** *Formation of Bends and Intrachain Cross-Linkages in Polypeptide Chains* In the following polypeptide, where might bends or turns occur? Where might intrachain disulfide cross-linkages be formed?

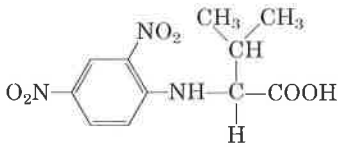
```
 1   2   3   4   5   6   7   8   9   10
Ile–Ala–His–Thr–Tyr–Gly–Pro–Phe–Glu–Ala–

 11  12  13  14  15  16  17  18  19  20
Ala–Met–Cys–Lys–Trp–Glu–Ala–Gln–Pro–Asp–

 21  22  23  24  25  26  27  28
Gly–Met–Glu–Cys–Ala–Phe–His–Arg–
```

**12.** *Location of Specific Amino Acids in Globular Proteins* X-ray analysis of the tertiary structure of myoglobin and other small, single-chain globular proteins has led to some generalizations about how the polypeptide chains of soluble proteins fold. With these generalizations in mind, indicate the probable location, whether in the interior or on the external surface, of the following amino acid residues in native globular proteins: Asp, Leu, Ser, Val, Gln, Lys. Explain your reasoning.

**13.** *The Number of Polypeptide Chains in an Oligomeric Protein* A sample (660 mg) of an oligomeric protein of $M_r$ 132,000 was treated with an excess of 1-fluoro-2,4-dinitrobenzene under slightly alkaline conditions until the chemical reaction was complete. The peptide bonds of the protein were then completely hydrolyzed by heating it with concentrated HCl. The hydrolysate was found to contain 5.5 mg of the following compound:




However, 2,4-dinitrophenyl derivatives of the $\alpha$-amino groups of other amino acids could not be found.

(a) Explain why this information can be used to determine the number of polypeptide chains in an oligomeric protein.

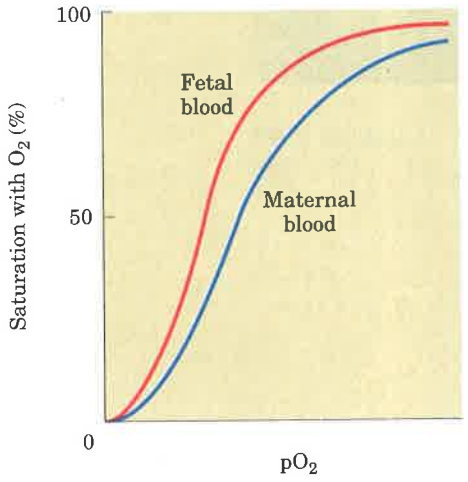
(b) Calculate the number of polypeptide chains in this protein.

**14.** *Molecular Weight of Hemoglobin* The first indication that proteins have molecular weights greatly exceeding those of the (then known) organic compounds was obtained over 100 years ago. For example, it was known at that time that hemoglobin contains 0.34% by weight of iron.

(a) From this information determine the minimum molecular weight of hemoglobin.

(b) Subsequent experiments indicated that the true molecular weight of hemoglobin is 64,500. What information did this provide about the number of iron atoms in hemoglobin?

**15.** *Comparison of Fetal and Maternal Hemoglobin* Studies of oxygen transport in pregnant mammals have shown that the $O_2$-saturation curves of fetal and maternal blood are markedly different when measured under the same conditions. Fetal erythrocytes contain a structural variant of hemoglobin, hemoglobin F, consisting of two $\alpha$ and two $\gamma$ subunits ($\alpha_2\gamma_2$), whereas maternal erythrocytes contain the usual hemoglobin A ($\alpha_2\beta_2$).




(a) Which hemoglobin has a higher affinity for oxygen under physiological conditions, hemoglobin A or hemoglobin F? Explain.

(b) What is the physiological significance of the different oxygen affinities? Explain.

CHAPTER

# Oxidative Phosphorylation and Photophosphorylation

Oxidative phosphorylation (ATP synthesis driven by electron transfer to oxygen) and photophosphorylation (ATP synthesis driven by light) are arguably the two most important energy transductions in the biosphere. These two processes together account for most of the ATP synthesized by aerobic organisms. Oxidative phosphorylation is the culmination of energy-yielding metabolism in aerobic organisms. All the enzymatic steps in the oxidative degradation of carbohydrates, fats, and amino acids in aerobic cells converge at this final stage of cellular respiration, in which electrons flow from catabolic intermediates to $O_2$, yielding energy for the generation of ATP from ADP and $P_i$. Photophosphorylation is the means by which photosynthetic organisms capture the energy of sunlight, the ultimate source of energy in the biosphere.

In eukaryotes, oxidative phosphorylation occurs in mitochondria; photophosphorylation occurs in chloroplasts. Oxidative phosphorylation involves the *reduction* of $O_2$ to $H_2O$ with electrons donated by NADH and $FADH_2$, and occurs equally well in light or darkness. Photophosphorylation involves the *oxidation* of $H_2O$ to $O_2$, with $NADP^+$ as electron acceptor, and it is absolutely dependent on light. These two highly efficient energy-conserving processes occur by fundamentally similar mechanisms.

Our current understanding of ATP synthesis in mitochondria and chloroplasts is based on a hypothesis, introduced by Peter Mitchell in 1961, in which transmembrane differences in proton concentration are central to energy transduction. This **chemiosmotic theory** has been accepted as one of the great unifying principles of twentieth century biology. It provides insight into the processes of oxidative phosphorylation and photophosphorylation, and into such apparently disparate energy transductions as active transport across membranes and the motion of bacterial flagella. Many biochemical details of these processes remain unsolved, but the chemiosmotic model described in this chapter provides the intellectual framework for investigating those details.

There are three fundamental similarities between oxidative phosphorylation and photophosphorylation. (1) Both processes involve the flow of electrons through a chain of redox intermediates, membrane-bound carriers that include quinones, cytochromes, and iron–sulfur proteins. (2) The free energy made available by this "downhill" (exergonic) electron flow is coupled to the "uphill" transport of protons across a proton-impermeable membrane, conserving some of the free energy of oxidation of metabolic fuels as a transmembrane electrochemical potential (p. 287). (3) The transmembrane flow of protons

down their concentration gradient through specific protein channels provides the free energy for synthesis of ATP.

In this chapter we first consider the process of oxidative phosphorylation. We begin with descriptions of the components of the electron transfer chain in mitochondria, the sequence in which these carriers act, and their organization into large functional complexes in the mitochondrial inner membrane. We then look at the chemiosmotic mechanism by which electron transfer is used to drive ATP synthesis, and the means by which this process is regulated in coordination with other energy-yielding pathways. The evolutionary origins of mitochondria, touched upon in Chapter 2, are further considered.

With this understanding of mitochondrial oxidative phosphorylation, we turn to photophosphorylation. Light-absorbing pigments in the membranes of chloroplasts and photosynthetic bacteria transfer the energy of absorbed light to reaction centers where electron flow is initiated. Electron flow occurs through a series of carriers and, as in mitochondria, this flow drives ATP synthesis by the chemiosmotic mechanism.

Albert L. Lehninger
1917–1986

## Mitochondrial Electron Flow

The discovery in 1948 by Eugene Kennedy and Albert Lehninger that mitochondria are the site of oxidative phosphorylation in eukaryotes marked the beginning of the modern phase of studies of biological energy transductions.

Mitochondria are organelles of eukaryotic cells, believed to have arisen during evolution when aerobic bacteria capable of oxidative phosphorylation took up symbiotic residence within a primitive, anaerobic, eukaryotic host cell (see Fig. 2–17). Mitochondria, like gram-negative bacteria, have two membranes (Fig. 18–1). The outer mitochondrial membrane is readily permeable to small molecules and ions; transmembrane channels composed of the protein porin allow most molecules of molecular weight less than 5,000 to pass easily. The inner membrane is impermeable to most small molecules and ions, including protons ($H^+$); the only species that cross the inner membrane are those for which there are specific transporter proteins. The inner membrane bears the components of the respiratory chain and the enzyme complex responsible for ATP synthesis.

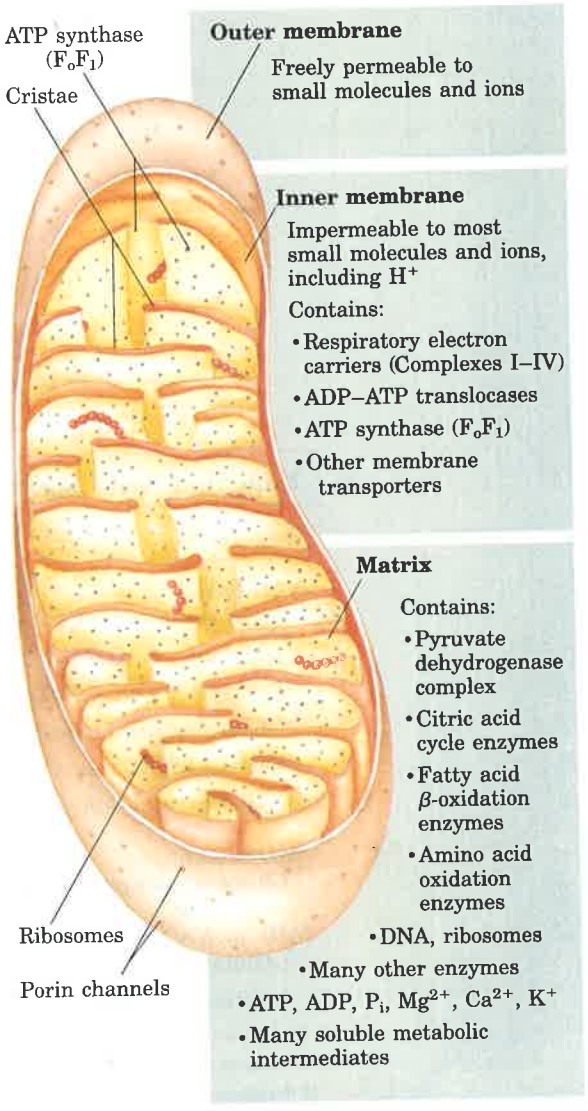



**Figure 18–1** Biochemical anatomy of a mitochondrion. The convolutions (cristae) of the inner membrane give it a very large surface area. The inner membrane of a single liver mitochondrion may have over 10,000 sets of electron transfer systems (respiratory chains) and ATP synthase molecules, distributed over the whole surface of the inner membrane. Heart mitochondria, which have very profuse cristae and thus a much larger area of inner membrane, contain over three times as many sets of electron transfer systems as liver mitochondria. The mitochondrial pool of coenzymes and intermediates is functionally separate from the cytosolic pool. The mitochondria of invertebrates, plants, and microbial eukaryotes are similar to those shown here, although there is much variation in size, shape, and degree of convolution of the inner membrane. See Chapter 2 for other details of mitochondrial structure.

Recall that the mitochondrial matrix, the space enclosed by the inner membrane, contains the pyruvate dehydrogenase complex and the enzymes of the citric acid cycle, the fatty acid β-oxidation pathway, and the pathways of amino acid oxidation—all of the pathways of fuel oxidation except glycolysis, which occurs in the cytosol. Because the inner membrane is selectively permeable, it segregates the intermediates and enzymes of cytosolic metabolic pathways from those of metabolic processes occurring in the matrix. Specific transporters carry pyruvate, fatty acids, and amino acids or their α-keto derivatives into the matrix for access to the machinery of the citric acid cycle. Similarly, ADP and $P_i$ are specifically transported into the matrix as the newly synthesized ATP is transported out.

We will discuss here in some detail the electron-carrying components of the mitochondrial respiratory chain.

## Electrons Are Funneled into Universal Electron Carriers

Most of the electrons entering the mitochondrial respiratory chain arise from the action of dehydrogenases that collect electrons from the oxidative reactions of the pyruvate dehydrogenase complex, the citric acid cycle, the β-oxidation pathway, and the oxidative steps of amino acid catabolism and funnel them as electron pairs into the respiratory chain. These dehydrogenases use either pyridine nucleotides (NAD or NADP; Table 18–1) or flavin nucleotides (FMN or FAD) as electron acceptors.

All of the pyridine nucleotide–linked dehydrogenases catalyze reversible reactions of the following general types:

$$\text{Reduced substrate} + NAD^+ \rightleftharpoons \text{oxidized substrate} + NADH + H^+$$

$$\text{Reduced substrate} + NADP^+ \rightleftharpoons \text{oxidized substrate} + NADPH + H^+$$

Most dehydrogenases are specific for $NAD^+$ as electron acceptor (Table 18–1), but some, such as glucose-6-phosphate dehydrogenase

**Table 18–1 Some important reactions catalyzed by NAD(P)H-linked dehydrogenases**

| Reaction* | Location† |
|---|---|
| *NAD-linked* | |
| α-Ketoglutarate + CoA + $NAD^+$ ⇌ succinyl-CoA + $CO_2$ + NADH + $H^+$ | M |
| L-Malate + $NAD^+$ ⇌ oxaloacetate + NADH + $H^+$ | M and C |
| Pyruvate + CoA + $NAD^+$ ⇌ acetyl-CoA + $CO_2$ + NADH + $H^+$ | M |
| Glyceraldehyde-3-phosphate + $P_i$ + $NAD^+$ ⇌ 1,3-bisphosphoglycerate + NADH + $H^+$ | C |
| Lactate + $NAD^+$ ⇌ pyruvate + NADH + $H^+$ | C |
| β-Hydroxyacyl-CoA + $NAD^+$ ⇌ β-ketoacyl-CoA + NADH + $H^+$ | M |
| *NADP-linked* | |
| Glucose-6-phosphate + $NADP^+$ ⇌ 6-phosphogluconate + NADPH + $H^+$ | C |
| *NAD- or NADP-linked* | |
| L-Glutamate + $H_2O$ + $NAD(P)^+$ ⇌ α-ketoglutarate + $NH_4^+$ + NAD(P)H | M |
| Isocitrate + $NAD(P)^+$ ⇌ α-ketoglutarate + $CO_2$ + NAD(P)H + $H^+$ | M and C |

* All of these reactions and their enzymes have been discussed in Chapters 14 through 17.

† M designates mitochondria; C, cytosol.

(see Fig. 14–22), require $NADP^+$. A few, such as glutamate dehydrogenase, can react with either $NAD^+$ or $NADP^+$. Some pyridine nucleotide–linked dehydrogenases are located in the cytosol, some in the mitochondria, and still others have two isozymes, one mitochondrial and the other cytosolic.

As was described in Chapter 13, the NAD-linked dehydrogenases remove two hydrogen atoms from their substrates. One of these is transferred as a hydride ion ($:H^-$) to the $NAD^+$; the other appears as $H^+$ in the medium (see Fig. 13–16). $NAD^+$ can also collect reducing equivalents from substrates acted upon by NADP-linked dehydrogenases. This is made possible by **pyridine nucleotide transhydrogenase,** which catalyzes the reaction

$$NADPH + NAD^+ \rightleftharpoons NADP^+ + NADH$$

NADH and NADPH are water-soluble electron carriers that associate *reversibly* with dehydrogenases. NADH acts as a diffusible carrier, transporting the electrons derived from catabolic reactions to their point of entry into the respiratory chain, the NADH dehydrogenase complex described below. NADPH is a diffusible carrier that supplies electrons to anabolic reactions.

Flavoproteins contain a very tightly, sometimes covalently, bound flavin nucleotide, either FMN or FAD (see Fig. 13–17). The oxidized flavin nucleotide can accept either one electron (yielding the semiquinone form) or two (yielding $FADH_2$ or $FMNH_2$). The standard reduction potential of a flavin nucleotide, unlike that of pyridine nucleotides, depends on the protein with which it is associated. Local interactions with functional groups in the protein distort the electron orbitals in the flavin ring, changing the relative stabilities of oxidized and reduced forms. The relevant standard reduction potential is therefore that of the particular flavoprotein, not that of isolated FAD or FMN, and the flavin nucleotide should be considered part of the flavoprotein's active site, not as a reactant or product in the electron transfer reaction. Because flavoproteins can participate in either one- or two-electron transfers, they can serve as intermediates between reactions in which two electrons are donated (as in dehydrogenations) and those in which only one electron is accepted (as in the reduction of a quinone to a hydroquinone, described below).

## Several Types of Electron Carriers Act in the Respiratory Chain

The mitochondrial respiratory chain consists of a series of electron carriers, most of which are integral membrane proteins, with prosthetic groups capable of accepting and donating either one or two electrons. Each component of the chain can accept electrons from the preceding carrier and transfer them to the following one, in a specific sequence. We noted earlier (Chapter 13) that there are four types of electron transfers in biological systems: (1) direct transfer of electrons, as in the reduction of $Fe^{3+}$ to $Fe^{2+}$; (2) transfer as a hydrogen atom ($H^+ + e^-$); (3) transfer as a hydride ion ($:H^-$), which bears two electrons; and (4) direct combination of an organic reductant with oxygen. Each of the first three types occurs in the respiratory chain. Some of the reactions in this sequence are one-electron transfers, and others involve the transfer of pairs of electrons. Whatever the form, the term reducing equivalent is used to designate a single electron equivalent that is transferred in an oxidation–reduction reaction.

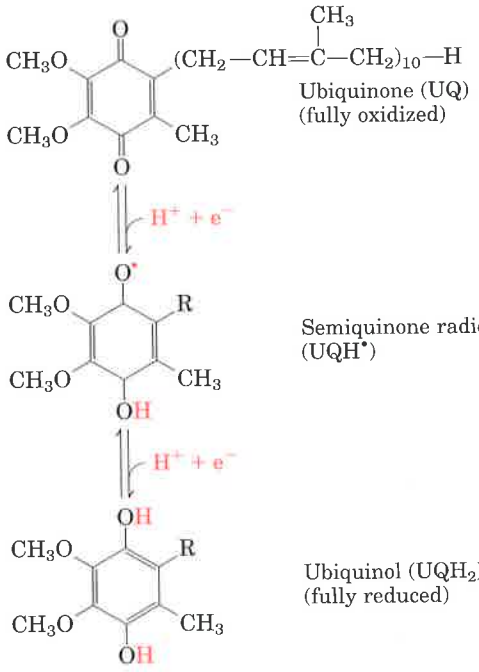


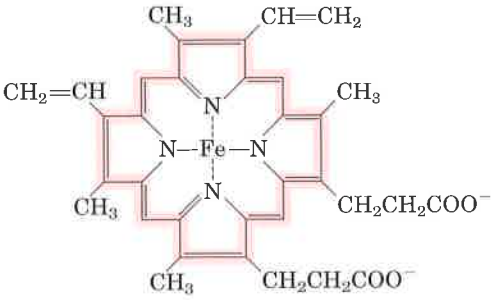

**Figure 18–2** Ubiquinone (UQ, or coenzyme Q), a respiratory chain electron carrier. Complete reduction of ubiquinone requires two electrons and two protons, and occurs in two steps through the semiquinone radical intermediate. The same chemistry is involved in the reduction of plastoquinone (a photosynthetic electron carrier) in chloroplasts and menaquinone (a respiratory chain carrier) in bacteria.

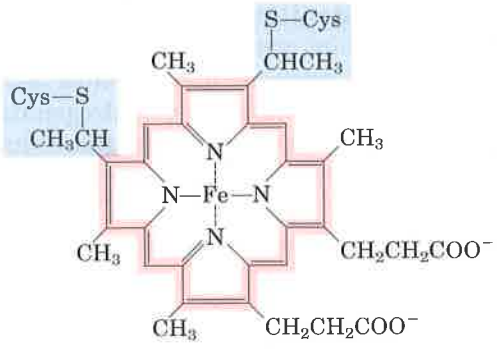
In addition to NAD and the flavoproteins described above, three other types of electron-carrying groups function in the respiratory chain: a hydrophobic benzoquinone (ubiquinone) and two different types of iron-containing proteins (cytochromes and iron–sulfur proteins).

**Ubiquinone** (also called **coenzyme Q,** or simply UQ) is a fat-soluble benzoquinone with a very long isoprenoid side chain (Fig. 18–2). The closely related compounds plastoquinone (found in plant chloroplasts) and menaquinone (found in bacteria) play roles analogous to that of ubiquinone; all carry electrons in membrane-associated electron transfer chains. Ubiquinone can accept one electron to become the semiquinone radical ($UQH^{\bullet}$) or two electrons to form ubiquinol ($UQH_2$) (Fig. 18–2) and, like flavoprotein carriers, it is therefore able to act at the junction between a two-electron donor and a one-electron acceptor. Because ubiquinone is both small and hydrophobic, it is freely diffusible within the lipid bilayer of the inner mitochondrial membrane, and can shuttle reducing equivalents between other, less mobile, electron carriers in the membrane.

The **cytochromes** are iron-containing electron transfer proteins of the mitochondrial inner membrane, the thylakoid membranes of chloroplasts, and the plasma membrane of bacteria. The characteristic strong colors of cytochromes are produced by the heme prosthetic group (Fig. 18–3).

There are three classes of cytochromes distinguished by differences in their light-absorption spectra and designated *a*, *b*, and *c*. Each type of cytochrome in its reduced ($Fe^{2+}$) state has three absorption bands in the visible range (Fig. 18–4). The longest-wavelength band (the $\alpha$ band) is near 600 nm in type *a* cytochromes, near 560 nm in type *b*, and near 550 in type *c*. To distinguish among closely related cyto-



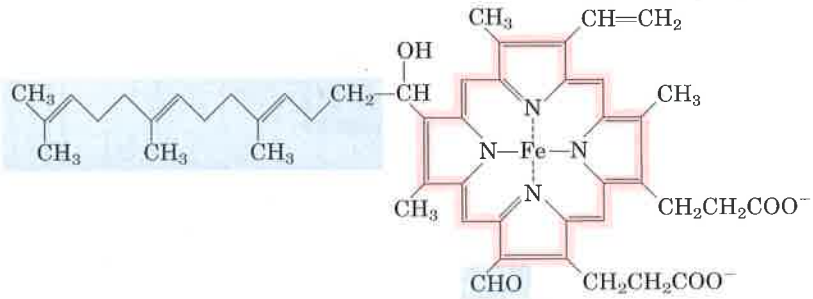




**Figure 18–3** The prosthetic groups of cytochromes have four five-membered, nitrogen-containing rings in a cyclic structure called a porphyrin. The four nitrogen atoms are coordinated with a central Fe ion that can be either $Fe^{2+}$ or $Fe^{3+}$. Iron protoporphyrin IX is found in *b*-type cytochromes and in hemoglobin and myoglobin (see Fig. 7–18). Heme C is bound covalently to the protein of cytochrome *c* through thioether bonds to two Cys residues. Heme A, found in the *a*-type cytochromes, has a long isoprenoid tail attached to one of the five-membered rings. The conjugated double bond system (shaded in red) of the porphyrin ring accounts for the absorption of visible light by these hemes.

chromes of one type, their exact absorption maximum is sometimes used in their names, as in cytochrome $b_{562}$ (the three-dimensional structure of this protein was shown in Fig. 7–25).

The heme groups of *a* and *b* cytochromes are tightly, but not covalently, bound to their associated proteins; heme groups of *c*-type cytochromes are covalently attached (through Cys residues; Fig. 18–3). As with the flavoproteins, the standard reduction potential of the iron atom in the heme of a cytochrome depends heavily on its interaction with protein side chains and is therefore different for each cytochrome. The cytochromes of type *a* and *b* and some of type *c* are integral membrane proteins. One striking exception is cytochrome *c* of mitochondria, a soluble protein that associates through electrostatic interactions with the outer surface of the mitochondrial inner membrane.

The ubiquitous occurrence of cytochrome *c* in aerobic organisms, together with its small size (104 amino acid residues), has allowed the determination of its amino acid sequence in many species from every phylum. The degree of sequence similarity in cytochrome *c* has been used as a measure of the evolutionary distances that separate species (see Fig. 6–16).

In some iron-containing electron transfer proteins, the **iron–sulfur proteins,** the iron is present not in heme (as it is in cytochromes) but in association with inorganic sulfur atoms and/or the sulfur atoms of Cys residues in the protein. These iron–sulfur (Fe–S) centers range from simple structures with a single Fe atom coordinated to four Cys residues through the sulfur in their side chains to more complex Fe–S centers with two or four Fe atoms (Fig. 18–5).

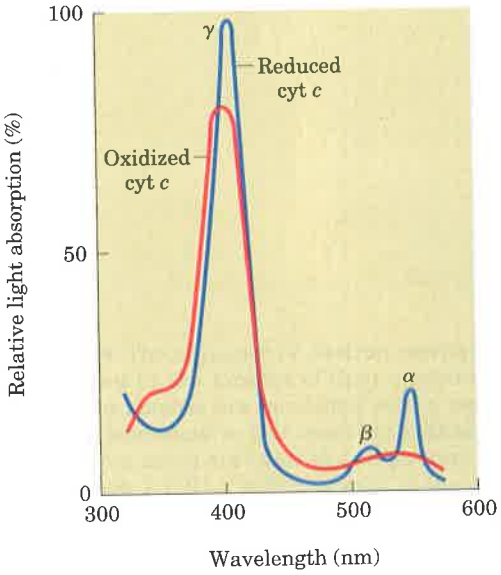



**Figure 18–4** Absorption spectra of cytochrome *c* in its oxidized (red) and reduced (blue) forms. Also labeled are the characteristic α, β, and γ bands of the reduced form.

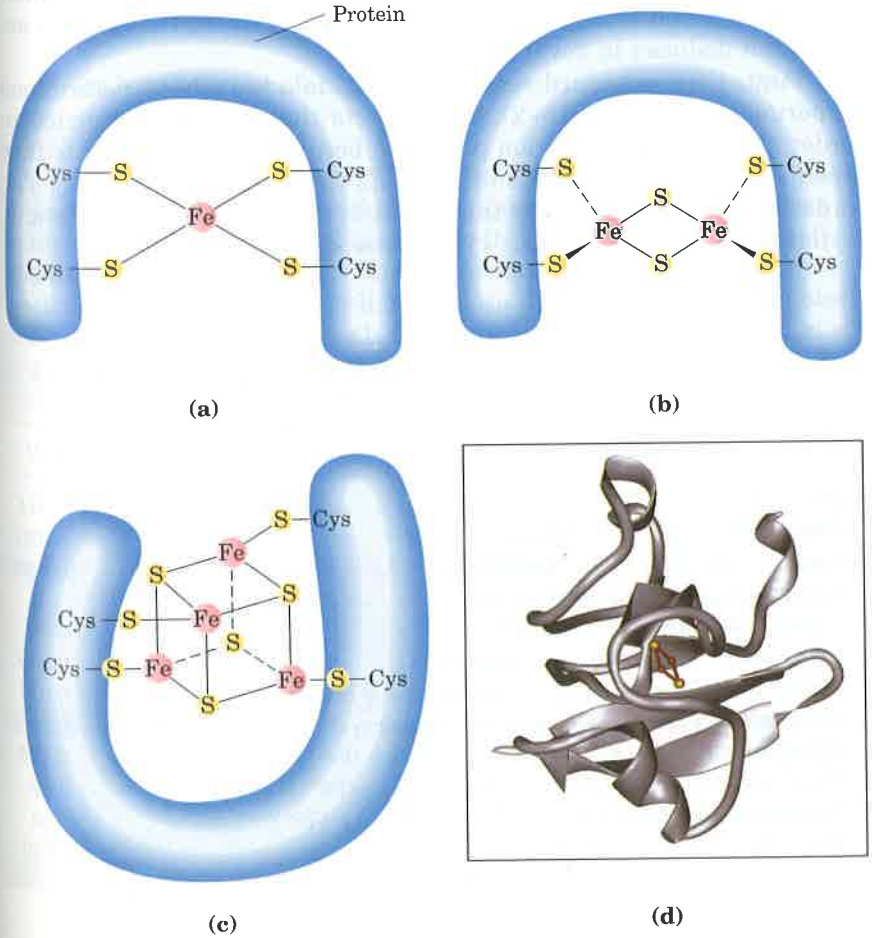



**Figure 18–5** The Fe–S centers of iron–sulfur proteins may be as simple as in **(a)**, with a single Fe ion surrounded by S atoms of four Cys residues. Other centers include both inorganic and Cys S atoms, as in the 2Fe–2S **(b)** or 4Fe–4S centers **(c)**. The ferredoxin in **(d)**, from the cyanobacterium *Anabaena* 7120, has one 2Fe–2S center. (Note that only the inorganic S atoms are counted in these designations. For example, in the 2Fe–2S center **(b)**, each Fe ion is actually surrounded by four S atoms.) The exact standard reduction potential of the iron in these centers depends on the type of center and the details of its interaction with its associated protein.

These proteins all participate in one-electron transfers, in which one of the Fe atoms is oxidized or reduced. With some exceptions, iron–sulfur proteins have low standard reduction potentials—they are good electron donors.

Many of the iron–sulfur proteins absorb visible light in the region of 400 to 460 nm, and in this region absorption decreases by about 50% when the proteins are reduced. Although the visible absorption spectrum can be used as a measure of oxidation state for purified proteins, this is not possible in more complex systems such as mitochondrial membranes, where the absorption is masked by the many other pigments present. However, the Fe atom(s) in iron–sulfur proteins are paramagnetic (i.e., they possess electrons with unpaired spins) and can therefore be detected by electron paramagnetic resonance (epr) spectroscopy. The epr signal, which is only observable at temperatures well below 0 °C, is the best measure of the presence, and the oxidation state, of a given iron–sulfur protein. The difficulty of detecting and identifying iron–sulfur proteins at room temperature has seriously complicated studies of their function in electron transfer, but it is clear that a number of iron–sulfur proteins play crucial roles in mitochondria (and chloroplasts).

## Mitochondrial Electron Carriers Function in Serially Ordered Complexes

In the overall reaction catalyzed by the mitochondrial respiratory chain, electrons move from NADH, succinate, or some other primary electron donor through flavoproteins, ubiquinone, iron–sulfur proteins, and cytochromes (nearly all of which are embedded in the inner membrane), and finally to $O_2$. The sequence in which the carriers act has been deduced in several ways.

First, their standard reduction potentials have been determined experimentally (Table 18–2). One expects the carriers to function in order of increasing reduction potential, because electrons tend to flow spontaneously from carriers of lower $E_0'$ to carriers of higher $E_0'$. The order of carriers deduced by this method is NADH, UQ, cytochrome *b*, cytochrome $c_1$, cytochrome *c*, cytochrome $a + a_3$.

**Table 18–2 Standard reduction potentials for respiratory chain and related electron carriers**

| Redox reaction (half-reaction) | $E_0'$ (V) |
|---|---|
| $2H^+ + 2e^- \longrightarrow H_2$ | −0.414 |
| $NAD^+ + H^+ + 2e^- \longrightarrow NADH$ | −0.320 |
| $NADP^+ + H^+ + 2e^- \longrightarrow NADPH$ | −0.324 |
| NADH dehydrogenase (FMN) + $2H^+ + 2e^- \longrightarrow$ NADH dehydrogenase ($FMNH_2$) | −0.30 |
| Ubiquinone + $2H^+ + 2e^- \longrightarrow$ ubiquinol | 0.045 |
| Cytochrome *b* ($Fe^{3+}$) + $e^- \longrightarrow$ cytochrome *b* ($Fe^{2+}$) | 0.077 |
| Cytochrome $c_1$ ($Fe^{3+}$) + $e^- \longrightarrow$ cytochrome $c_1$ ($Fe^{2+}$) | 0.22 |
| Cytochrome *c* ($Fe^{3+}$) + $e^- \longrightarrow$ cytochrome *c* ($Fe^{2+}$) | 0.254 |
| Cytochrome *a* ($Fe^{3+}$) + $e^- \longrightarrow$ cytochrome *a* ($Fe^{2+}$) | 0.29 |
| Cytochrome $a_3$ ($Fe^{3+}$) + $e^- \longrightarrow$ cytochrome $a_3$ ($Fe^{2+}$) | 0.55 |
| $\frac{1}{2}O_2 + 2H^+ + 2e^- \longrightarrow H_2O$ | 0.816 |

Second, when the entire chain of carriers is reduced experimentally by providing an electron source but no electron acceptor (no $O_2$), and then $O_2$ is suddenly introduced into the system, the rate at which each electron carrier becomes oxidized (measured spectroscopically) shows, in favorable cases, the order in which the carriers function (Fig. 18–6). The carrier nearest $O_2$ (at the end of the chain) gives up its electrons first, the second carrier from the end is oxidized next, and so on. Such experiments have confirmed the sequence deduced from standard reduction potentials.

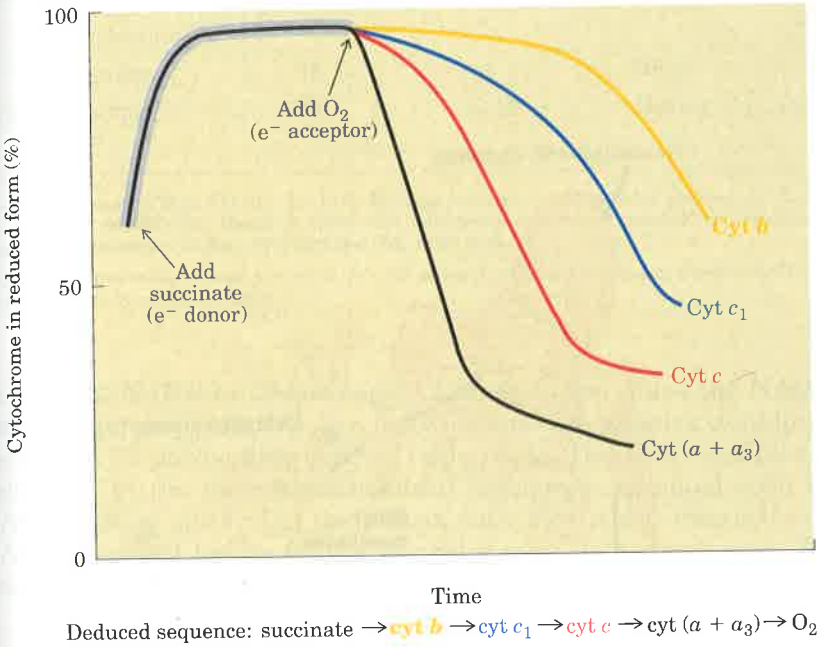



**Figure 18–6** The sequence of electron carriers can be determined by the kinetics of their oxidation. Isolated mitochondria are incubated with a source of electrons (succinate in this case) but without $O_2$. Electrons from succinate enter the respiratory chain (through $FADH_2$), reducing each of the electron carriers almost completely. Using rapid and sensitive spectrophotometric techniques, the rate of oxidation of each carrier is determined immediately after introducing $O_2$ into the system. The carriers closest to $O_2$ (cytochromes $a$ and $a_3$) are oxidized first; the most distant carrier (cytochrome $b$) is oxidized last.

Third, agents that inhibit the flow of electrons through the chain have been used in combination with measurement of the degree of oxidation of each carrier (Fig. 18–7). In the presence of $O_2$ and an electron donor, carriers that function before the inhibited step are expected to become fully reduced, and those that function after the block should be completely oxidized. By using several inhibitors that block earlier or later in the chain, the entire sequence has been deduced; it is the same as predicted from the first two approaches.

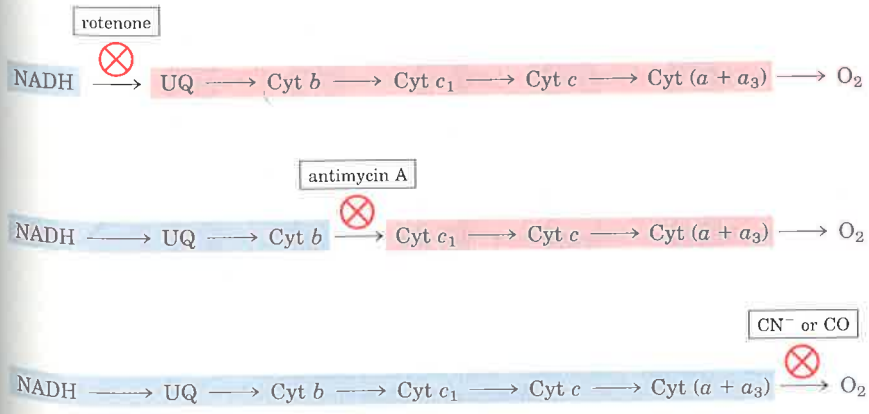



**Figure 18–7** Determination of the sequence of electron carriers by the effects of inhibitors of electron transfer on the oxidation state of each carrier. In the presence of an electron donor and $O_2$, each inhibitor causes a characteristic pattern of oxidized/reduced carriers: those before the block become reduced (blue), and those after the block become oxidized (red).

Fourth, gentle treatment of the inner mitochondrial membrane with detergents allows the resolution of four electron-carrier complexes, each representing a fraction of the entire respiratory chain (Fig. 18–8). Each of the four separated complexes has its own unique composition (Table 18–3), and each is capable of catalyzing electron transfer through a portion of the chain. Complexes I and II catalyze electron transfer to ubiquinone from two different electron donors: NADH (Complex I) and succinate (Complex II). Complex III carries electrons from ubiquinone to cytochrome *c*, and Complex IV completes the sequence by transferring electrons from cytochrome *c* to $O_2$ (Fig. 18–8).

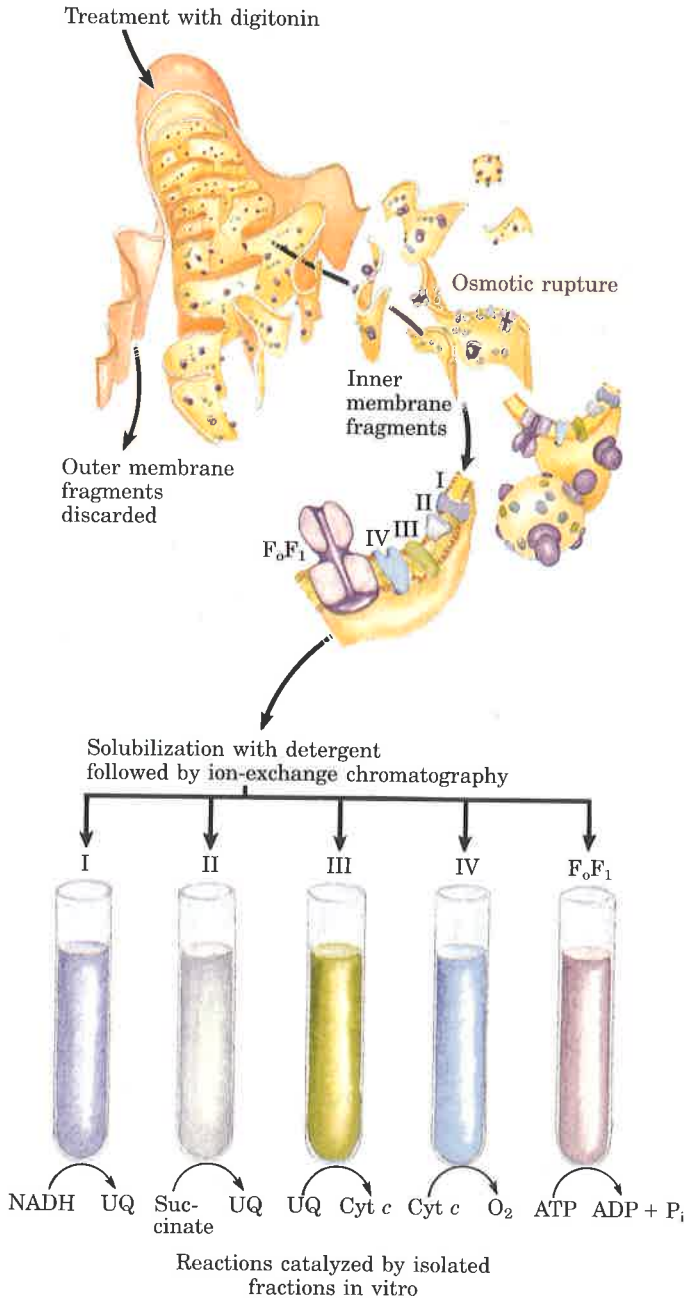



**Figure 18–8** Resolution of functional complexes of the respiratory chain. The outer mitochondrial membrane is first removed by treatment with the detergent digitonin. Fragments of inner membrane are then obtained by osmotic rupture of the mitochondria, and the fragments are gently dissolved in a second detergent. The resulting mixture of inner membrane proteins is resolved by ion-exchange chromatography into different complexes (I through IV) of the respiratory chain, each with its unique protein composition (see Table 18–3), and the enzyme ATP synthase (sometimes called Complex V). The isolated Complexes I through IV catalyze transfers between donors (NADH and succinate), intermediate carriers (UQ and cytochrome *c*), and $O_2$, as shown. ATP synthase, which is composed of peripheral ($F_1$) and integral ($F_o$) proteins (discussed later in this chapter), has only ATP-hydrolyzing (ATPase), not ATP-synthesizing, activity in vitro.

**Table 18–3 Protein components of the mitochondrial electron transfer chain**

| Enzyme complex* | Mass (kDa) | Number of subunits | Prosthetic group(s) |
|---|---|---|---|
| I NADH dehydrogenase | 850 | >25 | FMN, Fe–S |
| II Succinate dehydrogenase | 140 | 4 | FAD, Fe–S |
| III Ubiquinone-cytochrome *c* oxidoreductase | 250 | 10 | Hemes, Fe–S |
| Cytochrome *c* | 13 | 1 | Heme |
| IV Cytochrome oxidase | 160 | 6–13 | Hemes; $Cu_A$, $Cu_B$ |

*Sources:* DePierre, J.W. & Ernster, L. (1977) Enzyme topology of intracellular membranes. *Annu. Rev. Biochem.* **46,** 201–262; Hatefi, Y. (1985) The mitochondrial electron transport and oxidative phosphorylation system. *Annu. Rev. Biochem.* **54,** 1015–1069.

* Note that cytochrome *c* is not part of an enzyme complex, but moves between Complexes III and IV as a freely soluble protein.

***Complex I: NADH to Ubiquinone*** Complex I, also called the **NADH dehydrogenase complex,** is a huge flavoprotein complex containing more than 25 polypeptide chains (Table 18–3). The entire complex is embedded in the inner mitochondrial membrane, oriented with its NADH-binding site facing the matrix such that it can interact with NADH produced by any of the several matrix dehydrogenases. The overall reaction catalyzed by Complex I is

$$NADH + H^+ + UQ \rightleftharpoons NAD^+ + UQH_2$$

in which oxidized ubiquinone (UQ) accepts a hydride ion (two electrons and one proton) from NADH and a proton from the solvent water in the matrix. The enzyme complex first transfers a pair of reducing equivalents from NADH to its prosthetic group, FMN (Fig. 18–9). The complex also contains seven Fe–S centers of at least two different types, through which electrons pass on their way from FMN to ubiquinone. Amytal (a barbiturate drug), rotenone (a plant product commonly used as an insecticide), and the antibiotic piericidin A all inhibit electron flow from these Fe–S centers to ubiquinone (Table 18–4).

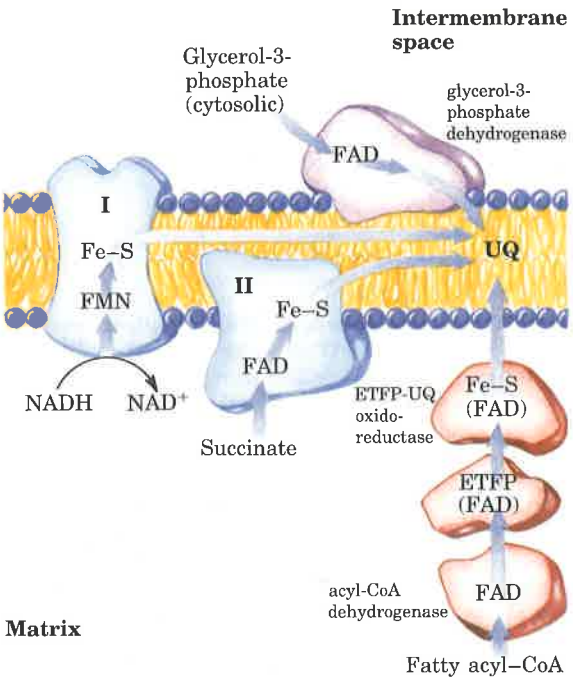



**Figure 18–9** Path of electrons from NADH, succinate, fatty acyl–CoA, and glycerol-3-phosphate to ubiquinone (UQ). Electrons from NADH pass through a flavoprotein to a series of iron–sulfur proteins (in Complex I) and then to UQ. Electrons from succinate pass through a flavoprotein and several Fe–S centers (in Complex II) on the way to UQ. Glycerol-3-phosphate donates electrons to a flavoprotein (glycerol-3-phosphate dehydrogenase) on the outer face of the inner mitochondrial membrane, from which they pass to UQ. Acyl-CoA dehydrogenase (the first enzyme in $\beta$ oxidation) transfers electrons to electron-transferring flavoprotein (ETFP), from which they pass via ETFP-ubiquinone oxidoreductase to UQ.

**Table 18–4 Some agents that interfere with oxidative phosphorylation or photophosphorylation**

| Type of interference | Compound | Target/mode of action |
|---|---|---|
| Inhibition of electron transfer | Cyanide, Carbon monoxide | Inhibit cytochrome oxidase |
| | Antimycin A | Blocks electron transfer from cytochrome *b* to cytochrome $c_1$ |
| | Rotenone, Amytal, Piericidin A | Prevent electron transfer from Fe–S center to ubiquinone |
| | DCMU | Competes with $Q_B$ for binding site in photosystem II |
| Inhibition of ATP synthase | Oligomycin, Venturicidin | Inhibit $F_1$ and $CF_1$ |
| | Dicyclohexylcarbodiimide (DCCD) | Blocks proton flow through $F_o$ and $CF_o$ |
| Uncoupling of phosphorylation from electron transfer | Carbonyl-cyanide phenylhydrazone, Dinitrophenol | Hydrophobic proton carriers |
| | Valinomycin | $K^+$ ionophore |
| | Uncoupling protein (thermogenin) | Forms proton-conducting pores in inner membrane of brown fat mitochondria |
| Inhibition of ATP–ADP exchange | Atractyloside | Inhibits adenine nucleotide translocase |

Ubiquinol ($UQH_2$, the fully reduced form; Fig. 18–2) diffuses in the membrane from Complex I to Complex III, where it is oxidized to UQ. The flow of electrons through Complex I to ubiquinone to Complex III is accompanied by the movement of protons from the mitochondrial matrix to the outer (cytosolic) side of the inner mitochondrial membrane (the intermembrane space), as described below.

***Complex II: Succinate to Ubiquinone*** We encountered Complex II in Chapter 15 under a different name: **succinate dehydrogenase;** it is the only membrane-bound enzyme in the citric acid cycle (p. 457). Although smaller and simpler than Complex I, it contains two types of prosthetic groups and at least four different proteins (Table 18–3). One protein has a covalently bound FAD and an Fe–S center with four Fe atoms; a second iron–sulfur protein is also present. Electrons are believed to pass from succinate to FAD, then through the Fe–S centers to ubiquinone.

Other substrates for mitochondrial dehydrogenases also pass electrons into the respiratory chain at the level of ubiquinone, but not through Complex II. The first step in the $\beta$ oxidation of fatty acyl–CoA, catalyzed by the flavoprotein **acyl-CoA dehydrogenase** (p. 486), involves transfer of electrons from the substrate to the FAD of the dehydrogenase, then to electron-transferring flavoprotein (ETFP), which in

turn passes its electrons to **ETFP-ubiquinone oxidoreductase** (Fig. 18–9). This reductase, an iron–sulfur protein that also contains a bound flavin nucleotide, passes electrons into the respiratory chain by reducing ubiquinone in the inner mitochondrial membrane.

In Chapter 16 we noted that glycerol released in the degradation of triacylglycerols is phosphorylated, then converted into dihydroxyacetone phosphate by **glycerol-3-phosphate dehydrogenase** (see Fig. 16–4). This enzyme is a flavoprotein located on the outer face of the inner mitochondrial membrane, and like succinate dehydrogenase and acyl-CoA dehydrogenase it channels electrons into the respiratory chain by reducing ubiquinone (Fig. 18–9). The important role of glycerol-3-phosphate dehydrogenase in shuttling reducing equivalents from cytosolic NADH into the mitochondrial matrix is described later (see Fig. 18–26).

***Complex III: Ubiquinone to Cytochrome* c** Complex III, also called **cytochrome $bc_1$ complex** or **ubiquinone-cytochrome *c* oxidoreductase,** contains cytochromes $b_{562}$ and $b_{566}$, cytochrome $c_1$, an iron–sulfur protein, and at least six other protein subunits (Table 18–3). These proteins are asymmetrically disposed in the inner mitochondrial membrane; cytochrome *b* spans the membrane, and both cytochrome $c_1$ and the iron–sulfur protein are on the outer surface. The switch between the two-electron carrier ubiquinone and the one-electron carriers (cytochromes $b_{562}$, $b_{566}$, $c_1$, and *c*) is accomplished in a series of reactions called the Q cycle (Fig. 18–10). Although the path of electron flow through this segment of the respiratory chain is complicated, the net effect of the transfer is simple: $UQH_2$ is oxidized to UQ and cytochrome *c* is reduced.

Complex III functions as a proton pump; as a result of the asymmetric orientation of the complex, protons produced when $UQH_2$ is oxidized to UQ are released to the intermembrane space, producing a transmembrane difference of proton concentration—a proton gradient. The importance of this proton gradient to mitochondrial ATP synthesis will soon become clear.

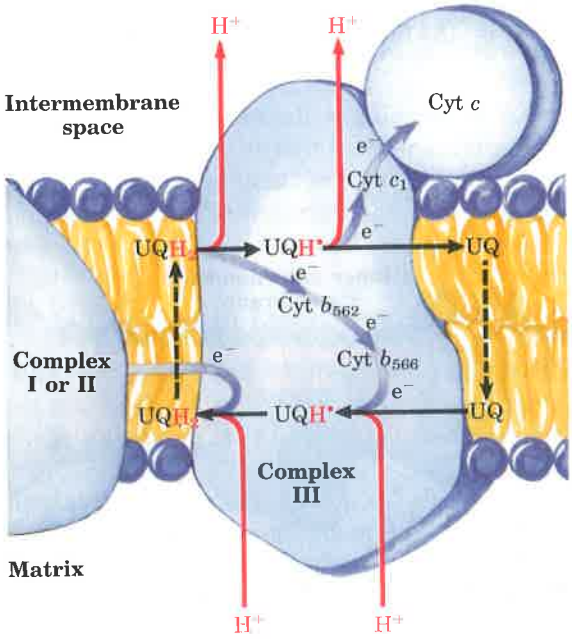



**Figure 18–10** The path of electrons through Complex III probably involves a "Q cycle" such as that shown here (blue arrows). The broken arrows represent diffusion of $UQH_2$ (ubiquinol) or its oxidized form UQ across the membrane. Notice that the electron transfers between cytochromes and ubiquinone are one-electron reactions, producing the semiquinone radical as an intermediate (see Fig. 18–2). The net effects of the reactions here are (1) movement of electrons from $UQH_2$ to cytochrome *c*, and (2) movement of protons from the inside (matrix) to the outer (cytosolic) side of the inner membrane (the intermembrane space). (For simplicity the oxidation–reduction cycles of the individual cytochromes are not shown here.)

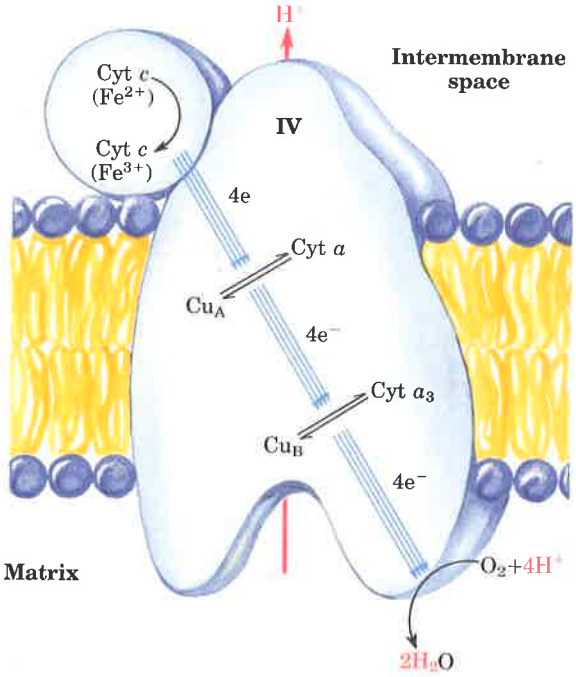



**Figure 18–11** Path of electrons through Complex IV. $Cu_A$ ($Cu^{2+}$) and cyt *a* ($Fe^{2+}$) form one bimetallic redox center capable of accepting two electrons; $Cu_B$ and cyt $a_3$ constitute a second two-electron redox center. The detailed path of electron flow between cyt *c* and $O_2$ is not known with certainty; apparently electrons first move from cyt *c* to $Cu_A$ or cyt *a*, which are in rapid redox equilibrium with each other. This bimetallic center then donates electrons to $Cu_B$ and cyt $a_3$, also in redox equilibrium, which in turn donate the electrons that reduce $O_2$ to $H_2O$. The four protons used in the reduction of $O_2$ to $H_2O$ are taken up from the matrix side of the inner mitochondrial membrane. Consequently, cytochrome oxidase pumps protons out of the matrix as electrons are transferred to $O_2$.

***Complex IV: Reduction of $O_2$*** Complex IV, also called **cytochrome oxidase,** contains cytochromes *a* and $a_3$. These cytochromes consist of two heme groups bound to different regions of the same large protein that are therefore spectrally and functionally distinct. Cytochrome oxidase also contains two copper ions, $Cu_A$ and $Cu_B$, that are crucial to the transfer of electrons to $O_2$. This complex enzyme has evolved to carry out the four-electron reduction of $O_2$ (Fig. 18–11) without generating incompletely reduced intermediates such as hydrogen peroxide or hydroxyl free radicals—very reactive species that would damage cellular components.

The flow of electrons from cytochrome *c* to $O_2$ through Complex IV causes net movement of protons from the matrix to the intermembrane space; Complex IV functions as a proton pump that contributes to the proton-motive force.

## Electron Transfer to $O_2$ Is Highly Exergonic

The energetics of electron transfer (oxidation–reduction) reactions were described in Chapter 13. In oxidative phosphorylation, two electrons pass from NADH through the respiratory chain to molecular oxygen:

$$NADH + H^+ + \tfrac{1}{2}O_2 \longrightarrow H_2O + NAD^+$$

For the redox pair $NAD^+$/NADH, $E'_0$ is −0.320 V, and for the pair $O_2/H_2O$, $E'_0$ is 0.816 V. The $\Delta E'_0$ for this reaction is therefore +1.14 V, and the standard free-energy change (p. 389) is

$$\begin{aligned}\Delta G^{\circ\prime} &= -n\mathcal{F}\Delta E'_0 \\ &= (-2)(96.5\ \text{kJ/V} \cdot \text{mol})(1.14\ \text{V}) \\ &= -220\ \text{kJ/mol}\end{aligned}$$

A similar calculation for succinate oxidation shows that electron transfer from succinate ($E'_0$ for fumarate/succinate = 0.031 V) to $O_2$ has a smaller, but still negative, standard free-energy change of about −152 kJ/mol.

In the mitochondrion, the combined action of Complexes I, III, and IV results in the transfer of electrons from NADH to $O_2$; Complexes II, III, and IV act together to catalyze electron transfer from succinate to $O_2$ (Fig. 18–12). The actual free-energy changes in respiring mitochon-

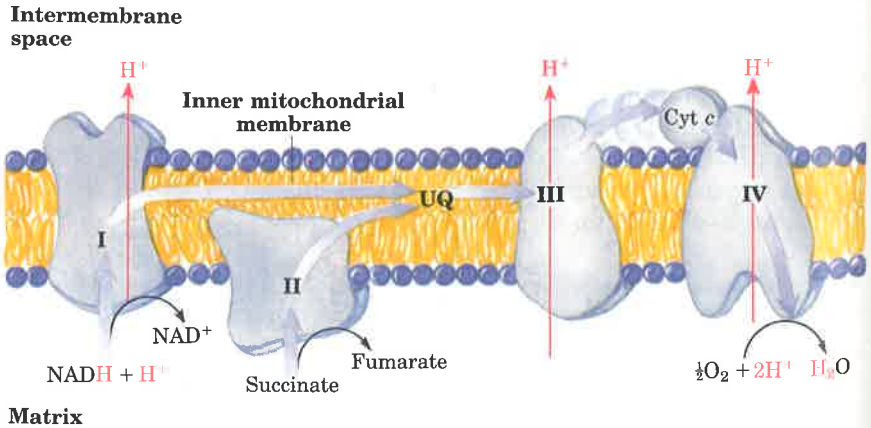



**Figure 18–12** Summary of the flow of electrons and protons through the four complexes of the respiratory chain. Electrons reach UQ via Complexes I and II. $UQH_2$ serves as a mobile carrier of electrons and protons. It passes electrons to Complex III, which passes them to another mobile connecting link, cytochrome *c*. Complex IV transfers electrons from reduced cytochrome *c* to $O_2$. Electron flow through Complexes I, III, and IV is accompanied by proton flow from the matrix to the intermembrane space. Recall that electrons from fatty acid β oxidation can also enter the respiratory chain through UQ (see Figs. 16–9 and 18–9).

dria are not the same as the standard free-energy changes, because reactants are not present at 1 M concentrations. However, it is clear from experimental measurements that under cellular conditions, the mitochondrial oxidation of NADH or succinate releases more free energy than the 51.8 kJ/mol required to drive the synthesis of ATP from ADP and $P_i$ (see Box 13–2).

## ATP Synthesis Coupled to Respiratory Electron Flow

We now turn to the most fundamental question about mitochondrial oxidative phosphorylation: how does the flow of electrons through the respiratory chain channel energy into the synthesis of ATP? We have seen that electron transfer through the respiratory chain releases more than enough free energy to form ATP. Mitochondrial oxidative phosphorylation therefore poses no thermodynamic problem. However, one cannot deduce from thermodynamic considerations the chemical mechanism by which energy released in one exergonic reaction (the oxidation of NADH by $O_2$) is channeled into a second, endergonic, reaction (the condensation of ADP and $P_i$). To describe the process of oxidative phosphorylation completely, we need to identify the physical and chemical changes that result from electron flow and cause ADP phosphorylation—the mechanism that *couples* oxidation with phosphorylation.

We begin our discussion by considering the stoichiometry of oxidation and phosphorylation in isolated mitochondria and the evidence for obligatory coupling of the two processes. The chemiosmotic interpretation of oxidative phosphorylation is then presented, with the major lines of evidence that support it. The enzyme ATP synthase, which is directly responsible for ATP synthesis, is the equivalent of an F-type ATP-dependent proton pump working in reverse; the flow of protons down their electrochemical gradient through this "pump" drives the condensation of $P_i$ and ADP. We describe also the membrane transport systems that move substrates, products, and reducing equivalents between the cytosol and the mitochondrial matrix. Having looked in detail at the coupling of ATP synthesis to electron flow, we will see that in the mitochondria of some tissues the two processes are deliberately "uncoupled" to produce heat.

We conclude with a summary of the overall regulation of ATP-producing processes in the cell, and a look at two further interesting aspects of mitochondria: the mitochondrial genome (and the effects of mutations therein) and the likely evolutionary origins of these organelles.

### Phosphorylation of ADP Is Coupled to Electron Transfer

When isolated mitochondria are suspended in a buffer containing ADP, $P_i$, and an oxidizable substrate such as succinate, three easily measured processes occur: (1) the substrate is oxidized (succinate yields fumarate), (2) $O_2$ is consumed (respiration occurs), and (3) ATP is synthesized. Careful experimental measurements of the stoichiometry of electron transfer to $O_2$ and the associated synthesis of ATP show that with NADH as electron donor, mitochondria synthesize nearly 3.0 ATP per pair of electrons passed to $O_2$, and with succinate nearly 2.0 ATP

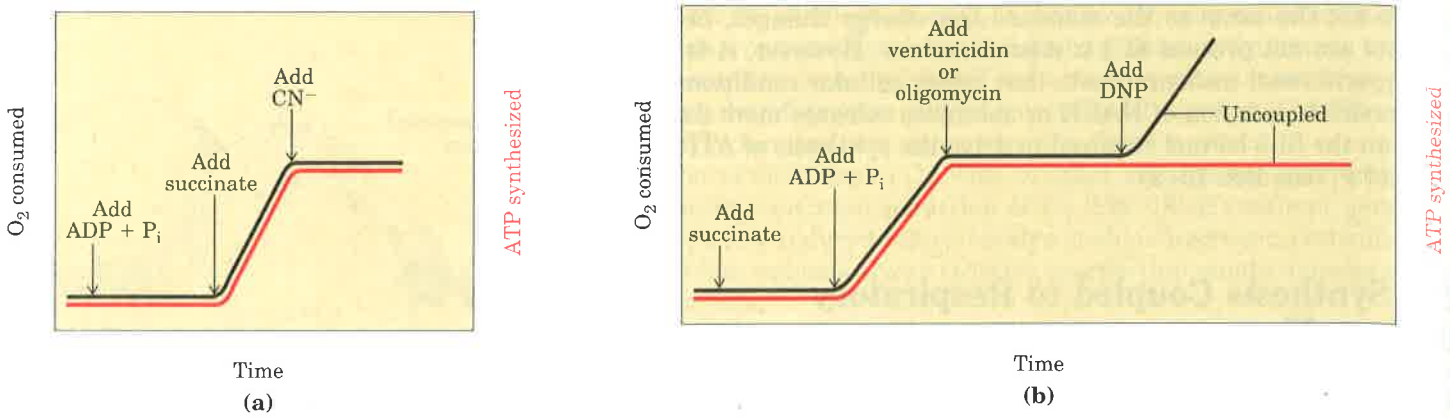
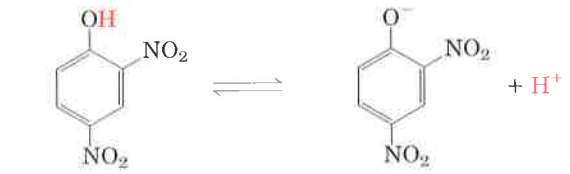
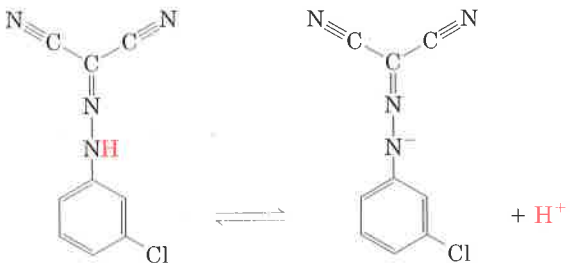



**Figure 18–13** Electron transfer to $O_2$ is tightly coupled to ATP synthesis in mitochondria, as is demonstrated in these experiments. Mitochondria are suspended in a buffered medium, and an $O_2$ electrode is used to monitor $O_2$ consumption. At intervals, samples are removed and assayed for the presence of ATP. **(a)** The addition of ADP and $P_i$ alone results in little or no increase in either respiration ($O_2$ consumption; black) or ATP synthesis (red). When succinate is added, respiration begins immediately and ATP is synthesized. The addition of cyanide ($CN^-$), which blocks electron transfer between cytochrome oxidase and $O_2$, inhibits both respiration and ATP synthesis. **(b)** Mitochondria provided with succinate respire and synthesize ATP only when ADP and $P_i$ are added. Subsequent addition of venturicidin or oligomycin, inhibitors of ATP synthase, blocks both ATP synthesis and respiration. Dinitrophenol (DNP) allows respiration to continue without ATP synthesis; DNP acts as an uncoupler.




2,4-Dinitrophenol (DNP)



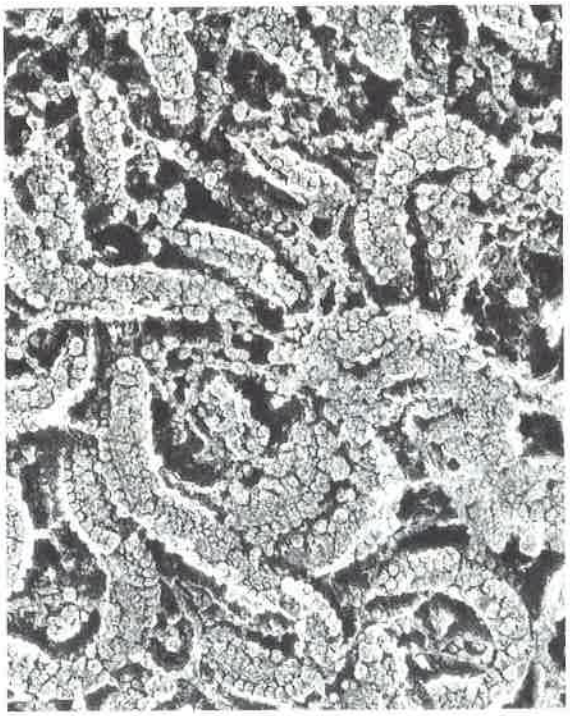

Carbonylcyanide-*m*-chlorophenylhydrazone (CCCP)

**Figure 18–14** Two chemical uncouplers of oxidative phosphorylation. Both have a dissociable proton and are very hydrophobic. They act by carrying protons across the inner mitochondrial membrane, dissipating the proton gradient. Both also uncouple photophosphorylation (p. 585).

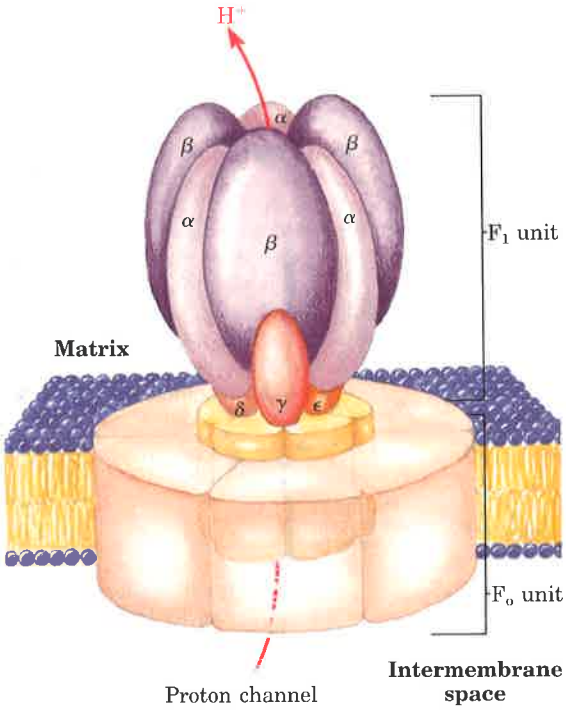
per electron pair. Oxygen consumption and ATP synthesis are dependent upon substrate oxidation, as can be seen in the experiments diagrammed in Figure 18–13.

Because the energy of substrate oxidation drives ATP synthesis in mitochondria, it is not surprising that inhibitors of the passage of electrons to $O_2$ (e.g., cyanide ion, carbon monoxide, and antimycin A) block ATP synthesis (Fig. 18–13a). It is perhaps not so obvious that the converse is true: inhibition of ATP synthesis blocks electron transfer in intact mitochondria. This obligatory coupling can be demonstrated in isolated mitochondria by providing $O_2$ and oxidizable substrates, but not ADP (Fig. 18–13b). Under these conditions, no ATP synthesis can occur, and electron transfer to $O_2$ is also strikingly reduced. Coupling of oxidation and phosphorylation can also be demonstrated using oligomycin or venturicidin, toxic antibiotics that bind to the ATP synthase in mitochondria. These compounds are potent inhibitors of both ATP synthesis *and* the transfer of electrons through the chain of carriers to $O_2$ (Fig. 18–13b). Because oligomycin is known not to interact directly with the electron carriers but only with ATP synthase, it follows that electron transfer and ATP synthesis are obligatorily coupled; neither reaction occurs without the other.

There are, however, certain conditions and reagents that uncouple oxidation from phosphorylation. When intact mitochondria are disrupted by treatment with detergent or physical shear, the resulting membrane fragments are still capable of catalyzing electron transfer from succinate or NADH to $O_2$, but no ATP synthesis is coupled to this respiration. Certain chemical compounds also cause uncoupling (Fig. 18–13b), without disrupting mitochondrial structure. The chemical uncouplers (Table 18–4) include 2,4-dinitrophenol (DNP) and a group of compounds related to carbonylcyanide phenylhydrazone (Fig. 18–14). All of these uncouplers are weak acids with hydrophobic properties. Ionophores (p. 560) also uncouple oxidative phosphorylation. These agents bind to inorganic ions and surround them with hydrophobic moieties; the ionophore–metal ion complexes pass easily through membranes. We shall see later how the chemiosmotic theory accounts for the action of uncouplers.



(a)




(b)

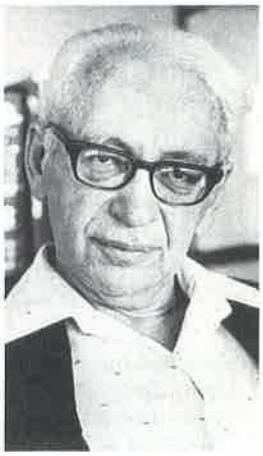
**Figure 18–15** The ATP synthase complex from mitochondria. **(a)** Electron micrographs showing the knoblike protrusions from the mitochondrial inner membrane. **(b)** Schematic diagram showing the likely organization of the subunits to form the proton-conducting $F_o$ portion, and the ATP-synthesizing $F_1$ unit. The $F_1$ complex consists of three $\alpha$, three $\beta$, and one each of $\gamma$, $\delta$, and $\epsilon$ subunits, as discussed later in the chapter (p. 562).

## ATP Synthase Is a Large Membrane–Protein Complex

**ATP synthase,** the ATP-synthesizing enzyme complex of the inner mitochondrial membrane, has two major components (or *f*actors), $F_1$ and $F_o$ (Fig. 18–15). The subscript letter *o* in $F_o$ denotes that it is the portion of the ATP synthase that confers sensitivity to *o*ligomycin, a potent inhibitor of this enzyme complex and thus of oxidative phosphorylation (Table 18–4).

$F_1$ was first extracted from the mitochondrial inner membrane and purified by Efraim Racker and his colleagues in the early 1960s. Isolated $F_1$ cannot synthesize ATP from ADP and $P_i$; because it can catalyze the reverse reaction—hydrolysis of ATP—the enzyme was originally called $F_1$ATPase. When $F_1$ is carefully extracted from inside-out vesicles prepared from the inner mitochondrial membrane (Fig. 18–16, p. 558), the vesicles still contain intact respiratory chains and can catalyze electron transfer, but cannot make ATP. When a preparation of isolated $F_1$ is added back to such depleted vesicles, their capacity to couple electron transfer and ATP synthesis is restored. Membrane reconstitution experiments of this kind, pioneered by Racker, opened new doors to research on membrane structure and function.

$F_1$, which in all aerobic organisms consists of six subunits, contains several binding sites for ATP and ADP, including the catalytic site for ATP synthesis. It is a peripheral membrane protein complex, held to the membrane by its interaction with $F_o$, an integral membrane protein complex of four different polypeptides that forms a transmembrane channel through which protons can cross the membrane. High-resolution electron micrographs of the isolated $F_oF_1$ complex show the knoblike $F_1$ head, a stalk, and a base piece ($F_o$), which normally extends across the inner membrane (Fig. 18–15b).



Efraim Racker
1913–1991

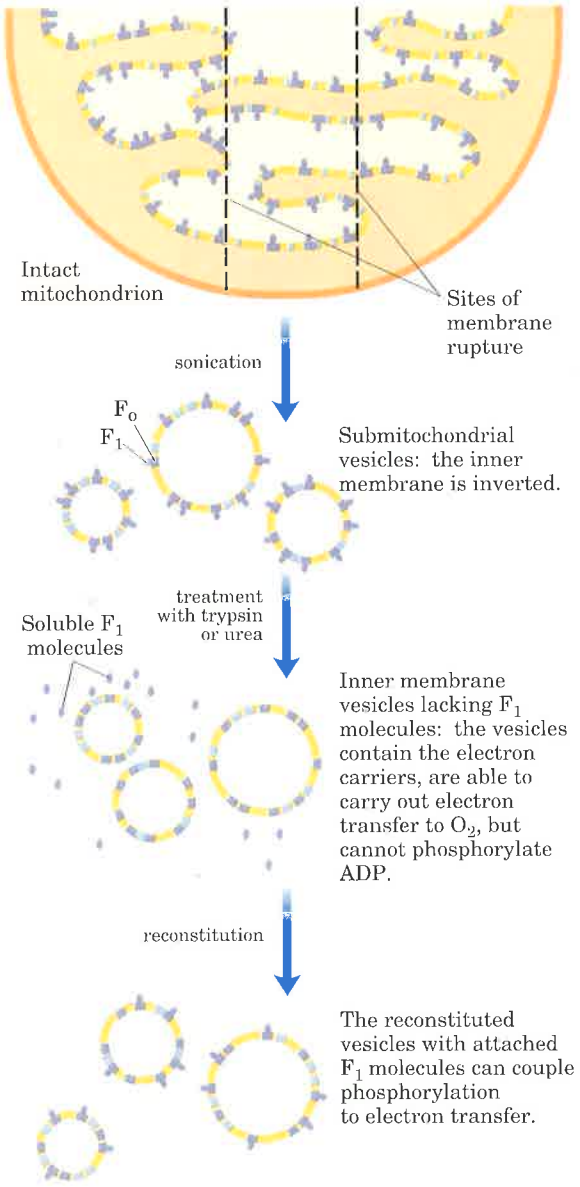
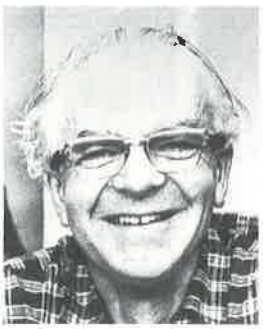



**Figure 18–16** Formation of inner-membrane vesicles by sonic treatment of mitochondria, and the reconstitution of oxidative phosphorylation by combining separated fractions of the membrane.



Peter Mitchell
1920–1992

The complete $F_oF_1$ complex, like isolated $F_1$, can hydrolyze ATP to ADP and $P_i$, but its biological function is to catalyze the condensation of ADP and $P_i$ to form ATP. The $F_oF_1$ complex is therefore more appropriately called ATP synthase.

### ATP Synthase Is Related to ATP-Dependent Proton Pumps

The structure and catalytic activity of ATP synthase from mitochondria show that it belongs to a larger class of enzymes that we have already encountered: the F-type ATPases (see Table 10–5). These protein complexes share the $F_oF_1$ structure, and there is sequence homology between the subunits of the F-type ATPases and those of mitochondrial ATP synthase. F-type ATPases are ATP-dependent proton pumps; they use the energy released by ATP hydrolysis to move protons across membranes against a concentration gradient, thereby creating a difference of pH and electrical charge (an electrochemical potential) across the membrane. The F-type ATPases must have evolved very early; they are found in the membranes of eubacteria and archaebacteria and in the mitochondria of all aerobic eukaryotes. Chloroplasts also contain an ATP synthase that is an F-type ATPase, which acts in light-driven ATP synthesis, as we shall see shortly.

### The Chemiosmotic Model: A Proton Gradient Couples Electron Flow and Phosphorylation

How does the oxidation of substrates via electron transfer through the respiratory chain cooperate with the ATP synthase to bring about phosphorylation of ADP to ATP? Early investigators of mitochondrial oxidative phosphorylation had before them a well-documented example of an oxidation coupled with a phosphorylation and involving a high-energy chemical intermediate: the glyceraldehyde-3-phosphate dehydrogenase reaction. In this glycolytic step, glyceraldehyde-3-phosphate is oxidized and simultaneously converted to 1,3-bisphosphoglycerate, a compound with a high-energy group at the site of the oxidation (see Fig. 14–5). ATP is formed when 1,3-bisphosphoglycerate transfers its activated $P_i$ to ADP. The formal similarity between this oxidative phosphorylation and that which takes place in mitochondria led naturally to the chemical coupling hypothesis, according to which some high-energy chemical intermediate is the direct product of electron transfer through one of three coupling sites along the chain of mitochondrial electron carriers. The energy in this putative chemical intermediate would then be used to drive the synthesis of ATP. An enormous amount of effort was expended in the search for such a chemical intermediate; none was found.

In the early 1960s Peter Mitchell suggested a new paradigm that has become central to current thinking and research on biological energy transductions. Mitchell's chemiosmotic hypothesis accounted for the experimental observations and made certain novel and experimentally testable predictions. Having survived more than two decades of vigorous testing and refinement, the hypothesis has become a generally accepted theory that is applied not only to mitochondrial oxidative phosphorylation, but to a wide array of other energy transductions, including light-driven ATP formation in photosynthetic organisms.

As postulated in the **chemiosmotic theory** (Fig. 18–17), the transfer of electrons along the respiratory chain is accompanied by outward pumping of protons across the inner mitochondrial mem-

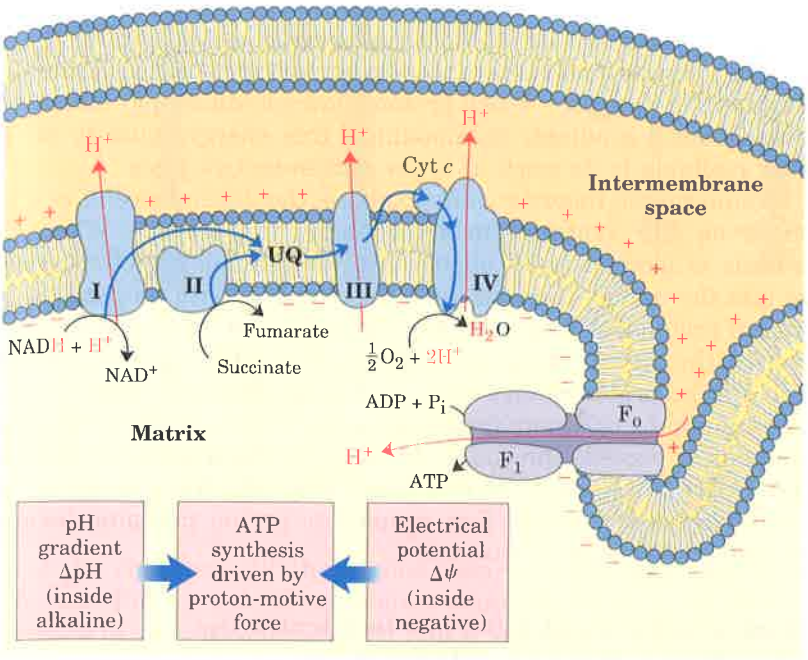



**Figure 18–17** In this simple version of the chemiosmotic theory applied to mitochondria, electrons from NADH and other oxidizable substrates pass through a chain of carriers (cytochromes, etc.) arranged asymmetrically in the membrane. Electron flow is accompanied by proton transfer across the mitochondrial membrane, producing both a chemical (ΔpH) and an electrical ($\Delta\psi$) gradient. The inner mitochondrial membrane is impermeable to protons; protons can reenter the matrix only through proton-specific channels ($F_o$). The proton-motive force that drives protons back into the matrix provides the energy for ATP synthesis, catalyzed by the $F_1$ complex associated with $F_o$.

brane, which results in a transmembrane difference in proton concentration (a proton gradient) and thus in pH; the matrix becomes alkaline relative to the cytosolic side of the membrane. The electrochemical energy inherent in this difference in proton concentration and separation of charge, the **proton-motive force,** represents a conservation of part of the energy of oxidation. The proton-motive force is subsequently used to drive the synthesis of ATP catalyzed by $F_1$ as protons flow passively back into the matrix through proton pores formed by $F_o$. To emphasize this crucial role of the proton-motive force, the equation for ATP synthesis is sometimes written

$$ADP + P_i + \{H^+\}_{out} \rightleftharpoons ATP + H_2O + \{H^+\}_{in} \qquad (18\text{–}1)$$

In the more general case, the electrochemical energy of a transmembrane gradient of any charged species is seen to be interconvertible with the energy of chemical bonds. The energy stored in such a gradient has two components, as implied in the word "electrochemical." One is the chemical potential energy due to the difference in concentration of a chemical species in the two regions separated by the membrane; the other is the electrical potential energy that results from the separation of charge when an ion moves across the membrane without its counterion (Fig. 18–17).

We showed in Chapter 10 that the free-energy change for the creation of an electrochemical gradient by an ion pump (Fig. 18–18) is

$$\Delta G = RT \ln (C_2/C_1) + Z\mathcal{F}\Delta\psi \qquad (18\text{–}2)$$

where $C_2/C_1$ is the concentration ratio for the ion that moves, $Z$ is the absolute value of its electrical charge (1 for a proton), and $\Delta\psi$ is the transmembrane difference in electrical potential, measured in volts.

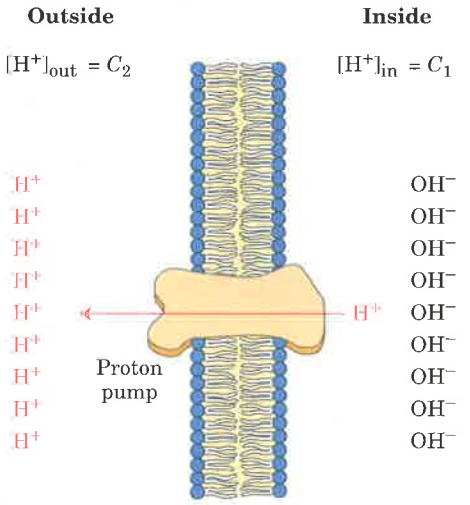
For the case of protons at 25 °C

$$\ln (C_2/C_1) = 2.3(\log [H^+]_{out} - \log [H^+]_{in}) = 2.3(pH_{in} - pH_{out}) = 2.3\ \Delta pH$$

and Equation 18–2 reduces to

$$\Delta G = 2.3RT\ \Delta pH + \mathcal{F}\Delta\psi$$

$$\Delta G = (5.70\ kJ/mol)\Delta pH + (96.5\ kJ/V \cdot mol)\Delta\psi \qquad (18\text{–}3)$$




$$\Delta G = RT \ln (C_2/C_1) + Z\mathcal{F}\Delta\psi$$
$$= 2.3RT\ \Delta pH + \mathcal{F}\Delta\psi$$

**Figure 18–18** The inner mitochondrial membrane separates two compartments of different pH, producing differences in both chemical concentration (ΔpH) and charge distribution, creating an electrical potential difference ($\Delta\psi$). The net effect of these differences is the proton-motive force ($\Delta G$), which can be calculated as shown. This is explained more fully in the text.

For the pumping of ions against an electrochemical gradient, both terms on the right-hand side of this equation have positive values, and $\Delta G$ is therefore positive. When protons flow spontaneously *down* their electrochemical gradient, an amount of free energy equal to $\Delta G$ becomes available to do work; it is the proton-motive force.

Chemiosmotic theory readily explains the dependence of electron transfer on ATP synthesis in mitochondria (Fig. 18–17). When ATP synthesis is blocked (with oligomycin, for example), protons cannot flow into the matrix through the $F_oF_1$ complex. With no path for the return of protons to the matrix and the continued extrusion of protons driven by the activity of the respiratory chain, a large proton gradient, hence a large proton-motive force, builds up. When the cost (free energy) of pumping one more proton out of the matrix against this gradient equals or exceeds the energy released by the transfer of electrons from NADH to $O_2$, electron flow must stop; the free energy for the overall process of electron flow coupled to proton pumping becomes zero, and equilibrium is attained.

## Strong Experimental Evidence Implicates the Proton-Motive Force in ATP Synthesis

The transmembrane proton gradient predicted by chemiosmotic theory has been experimentally measured. When intact mitochondria are suspended with an oxidizable substrate such as succinate in a lightly buffered medium, the addition of $O_2$ results in acidification of the suspending medium (Fig. 18–19). The stoichiometry of proton pumping can be estimated from the pH changes, after correcting for buffering effects. The measurement is difficult, and there is no universal agreement on the stoichiometry; about ten protons are pumped out for each pair of electrons transferred from NADH to $O_2$. The result is a steady state with a difference of about one pH unit between the matrix and the medium, inside (matrix) alkaline. Indirect measurements of the transmembrane electrical potential in respiring mitochondria yield a value of 0.1 to 0.2 V, inside negative. Substituting these values into the equation for proton-motive force (Eqn 18–3) gives a value of 15 to 25 kJ/mol, the free-energy change for the transmembrane movement of one equivalent of protons back into the mitochondrial matrix. This free-energy change is large enough to account for the synthesis of one ATP when two to three protons flow into the matrix.

Uncouplers are hydrophobic weak acids (Fig. 18–14). Their hydrophobicity allows them to diffuse readily across mitochondrial membranes. After entering the mitochondrial matrix in the protonated form, they can release a proton (dissociate), thus dissipating the proton gradient. **Ionophores** uncouple electron transfer from oxidative phosphorylation by creating electrical short circuits across the mitochondrial membrane (Fig. 18–20). The toxic ionophore valinomycin forms a lipid-soluble complex with $K^+$, which is abundant in the cytosol (see Fig. 10–26). Whereas $K^+$ penetrates the mitochondrial inner membrane only very slowly, the $K^+$–valinomycin complex readily passes through. The influx of positive ions neutralizes the excess of negative charge inside the matrix, diminishing the electrical component of the proton-motive force. Valinomycin slows mitochondrial ATP synthesis without blocking electron transfer to $O_2$ (Table 18–4).

If the role of electron transfer in mitochondrial ATP synthesis is simply to pump protons to create the electrochemical potential of the proton-motive force, an artificially created proton gradient should be able to replace electron transfer in driving ATP synthesis. This predic-

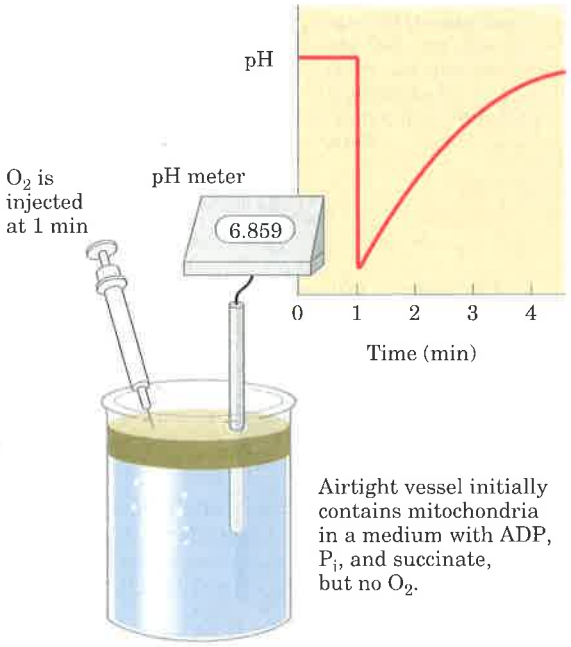



**Figure 18–19** Isolated mitochondria are suspended in a medium containing ADP, $P_i$, and succinate, but initially no $O_2$. When a small amount of $O_2$ is injected into the reaction mixture, succinate oxidation and electron transfer to $O_2$ begin immediately. A pH electrode registers a sudden decrease in the pH of the medium, indicating that protons are moving out of the mitochondria. As the injected $O_2$ is consumed, protons slowly leak back into the mitochondria, and the external pH returns to the initial level. From the known amount of $O_2$ added and the measured pH change, one can in principle calculate the number of protons extruded per molecule of $O_2$ consumed.

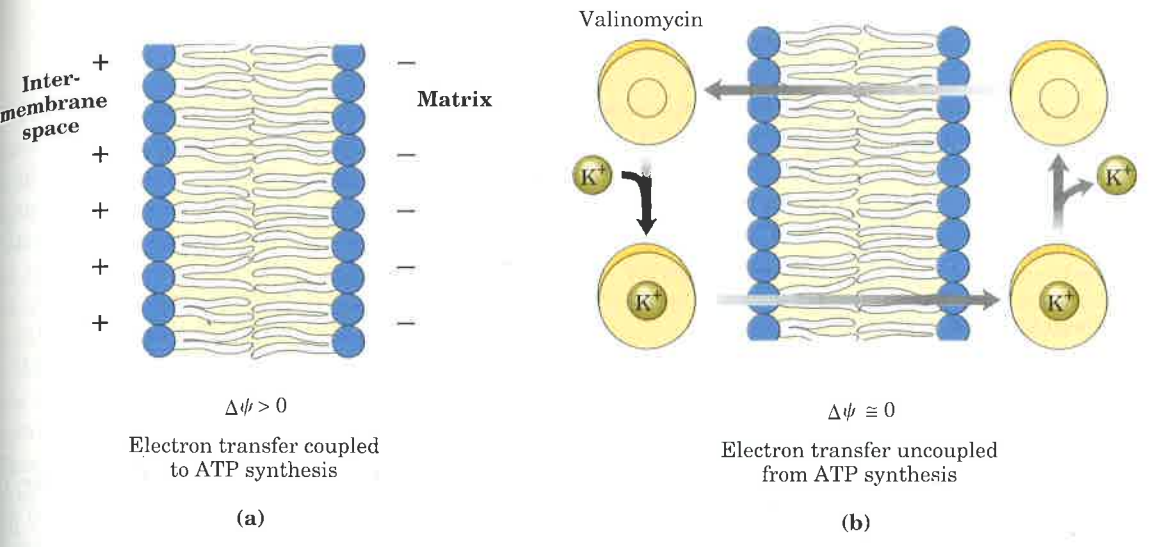



**Figure 18–20** Ionophores such as valinomycin uncouple oxidative phosphorylation by dissipating ion gradients across the inner mitochondrial membrane, eliminating the contribution of $\Delta\psi$ to the proton-motive force. **(a)** A transmembrane electrical potential ($\Delta\psi$) exists because of the unequal distribution of protons on either side. **(b)** Valinomycin combines reversibly with $K^+$ ions to form a membrane-permeable complex that diffuses across the inner membrane and releases $K^+$ on the inside. This movement of charge reduces the value of $\Delta\psi$, and uncouples electron transfer from ATP synthesis.

n of the chemiosmotic theory has been experimentally tested and nfirmed (Fig. 18–21). Mitochondria manipulated so as to impose a fference of proton concentration and a separation of charge across e inner membrane (Fig. 18–21) carry out the synthesis of ATP *in the sence of an oxidizable substrate;* the proton-motive force alone suf- es to drive ATP synthesis.

Chemiosmotic theory also accounts for a third condition that un- uples oxidation from phosphorylation—mechanical disruption of the itochondrial membrane. Without an intact membrane there can be proton gradient, hence no energy conservation and no ATP synthe- s.

## he Detailed Mechanism of ATP Formation Remains Elusive

though it is clear that a transmembrane proton gradient provides e energy for ATP synthesis, it is not clear how this energy is trans- itted to the ATP synthase. This enzyme catalyzes the conversion of zyme-bound ADP and $P_i$ into bound ATP even in the absence of a oton gradient. Remarkably, it appears that the reaction ADP + $\rightleftharpoons$ ATP + $H_2O$ is readily reversible—that the free-energy change

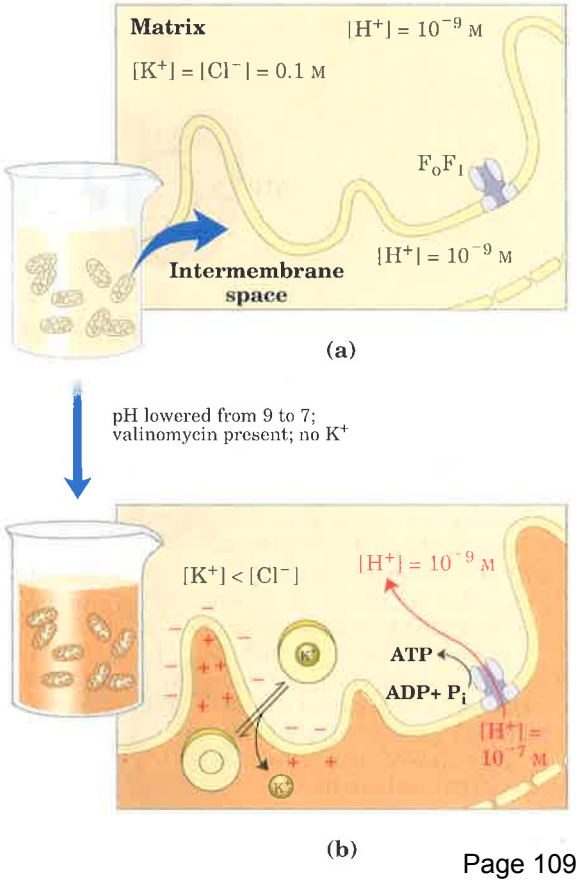



**Figure 18–21** An artificially imposed electrochemical gradient can drive ATP synthesis in the absence of an oxidizable substrate as electron donor. In this two-step experiment, isolated mitochondria are first incubated in a pH 9 buffer containing 0.1 M KCl **(a)**. The slow leakage of buffer and KCl into the mitochondria eventually brings the matrix into equilibrium with the surrounding medium. No oxidizable substrates are present. **(b)** In the second step, mitochondria are separated from the pH 9 buffer and resuspended in pH 7 buffer containing valinomycin but no KCl. The change in buffer creates a difference of two pH units. The outward flow of $K^+$, carried (without a counterion) down its concentration gradient by valinomycin, creates a charge imbalance across the membrane (matrix negative). The sum of the chemical potential provided by the pH difference and the electrical potential provided by the separation of charges is a proton-motive force large enough to support ATP synthesis in the absence of an oxidizable substrate.

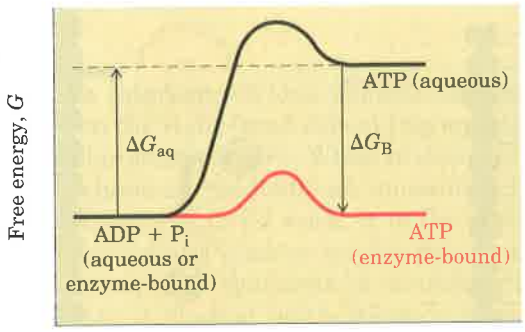



**Figure 18–22** Reaction coordinate diagram for the condensation of ADP and $P_i$ to form ATP on the surface of ATP synthase. Although the free-energy change for this reaction in aqueous solution ($\Delta G_{aq}$) is large and positive, the very tight binding of ATP to the enzyme provides binding energy ($\Delta G_B$) that brings the free energy of enzyme-bound ATP close to that of ADP + $P_i$. On the enzyme surface, the reaction is therefore readily reversible; the equilibrium constant is believed to be near 1.

for ATP synthesis on the enzyme surface is close to zero. However, the ATP formed in this reaction remains tightly bound to the active site, preventing further catalysis at that site. The proton-motive force apparently supplies the energy needed to force dissociation of the tightly bound ATP from the enzyme, allowing another catalytic cycle to begin.

How is it possible that ATP synthesis, known to be strongly endergonic in aqueous solution ($\Delta G^{\circ\prime}$ = 30.5 kJ/mol), is readily reversible on the enzyme surface? The very tight noncovalent binding of enzyme to ATP may supply enough binding energy (Chapter 8) to make bound ATP about as stable as its hydrolysis products (Fig. 18–22). Alternatively, the reaction may occur in a very hydrophobic pocket in the enzyme interior, where the energetics of hydrolysis in water do not apply.

Each $F_1$ complex has the subunit composition $\alpha_3\beta_3\gamma\delta\epsilon$ (Fig. 18–15). A tight binding site for ATP, apparently identical to the catalytic site, is located on each $\beta$ subunit, or perhaps between each $\beta$ and its associated $\alpha$ subunit. Every $F_1$ complex therefore has three ATP-synthesizing sites, which interact with $F_o$ through the single copies of $\gamma$, $\delta$, and $\epsilon$ subunits.

On the basis of detailed kinetic and binding studies of the reactions catalyzed by $F_oF_1$, Paul Boyer has suggested a mechanism (Fig. 18–23) in which the three active sites on $F_1$ alternate in catalyzing ATP synthesis. The limiting step in the process is the release of newly synthesized ATP from the enzyme. A conformational transition driven by the proton-motive force reduces the enzyme's affinity for ATP.

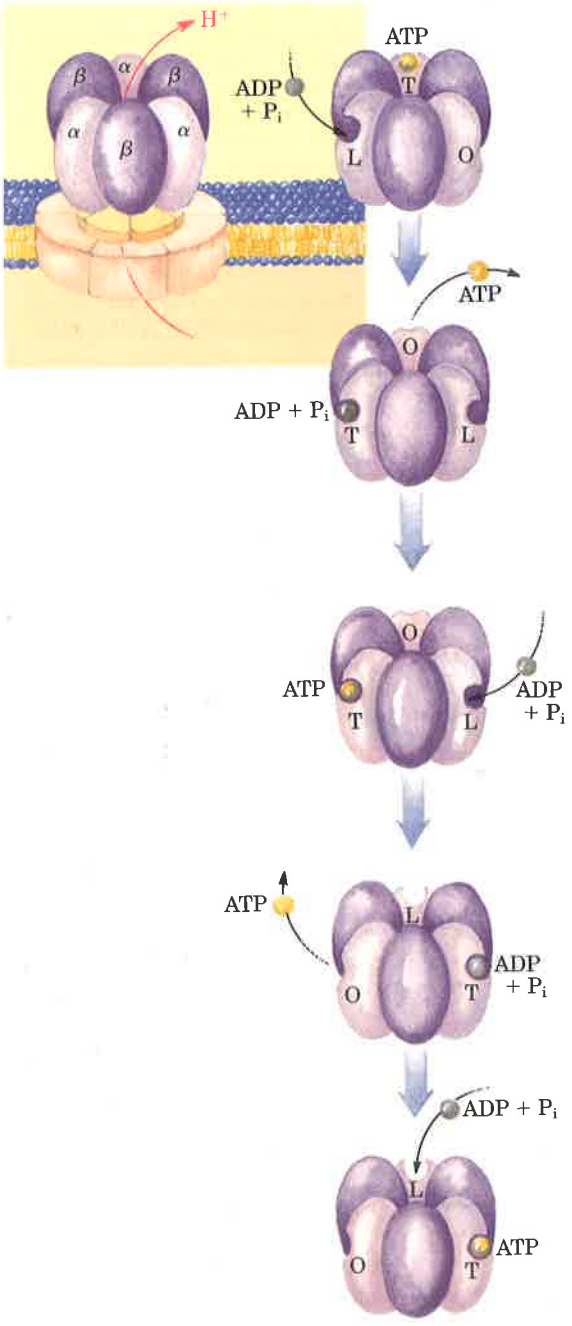



**Figure 18–23** The "binding change" model for ATP synthase action. The enzyme has three equivalent adenine nucleotide binding sites, one for each pair of $\alpha$ and $\beta$ subunits. At any given moment, one of these sites is in the T (tight-binding) conformation, a second is in the L (loose-binding) conformation, and a third is in the O (open; very loose-binding) conformation. At the beginning of a catalytic cycle, the T site is occupied by ATP, and ADP and $P_i$ bind loosely to the L site. The proton-motive force causes, perhaps by the flow of protons through the $F_o$ channel, a cooperative conformational change in which the T site is converted to an O site, and ATP dissociates from it; the L site is converted to a T site, where ADP and $P_i$ quickly condense to form ATP; and the O site becomes an L site, where ADP and $P_i$ loosely bind. The experimental results require that at least two of the three catalytic sites alternate in activity; ATP cannot be released from one site unless and until ADP and $P_i$ are bound at the other.

The transition induced by the proton-motive force may be envisioned as a 120° rotation of the $\alpha_3\beta_3$ portion of $F_1$ (Fig. 18–23), placing one of the three $\alpha$–$\beta$ pairs in a special position relative to the proton channel of $F_o$. However, no physical rotation of $F_1$ has been demonstrated, and the model does not require it; the conformational change that interconverts the three types of sites may be an allosteric transition, in which changes at one $\alpha$–$\beta$ pair force compensating changes in the other two pairs.

What makes the mechanism of this enzyme particularly difficult to solve is the *vectorial* nature of the process it catalyzes. Somehow, the enzyme must sense not merely a certain concentration of protons, but a difference of proton concentration in two regions of space. Determination of the detailed structure of $F_1$ by x-ray crystallography should shed light on the mechanism of its action.

## The Proton-Motive Force Energizes Active Transport

The primary role of electron transfer in mitochondria is to furnish energy for the synthesis of ATP during oxidative phosphorylation, but this energy serves also to drive several transport processes essential to oxidative phosphorylation. We have seen that the inner mitochondrial membrane is generally impermeable to charged species, but two specific systems in the inner mitochondrial membrane transport ADP and $P_i$ into the matrix and ATP out to the cytosol (Fig. 18–24). The **adenine nucleotide translocase,** which extends across the inner membrane, binds $ADP^{3-}$ on the outside (cytosolic) surface of the inner membrane and transports it inward in exchange for an $ATP^{4-}$ molecule simultaneously transported outward (see Fig. 13–1 for the ionic forms of ATP and ADP). Because this antiporter moves four negative charges out for each three moved in, its activity is favored by the transmembrane electrochemical gradient, which gives the matrix a net negative charge; the proton-motive force drives ATP–ADP exchange.

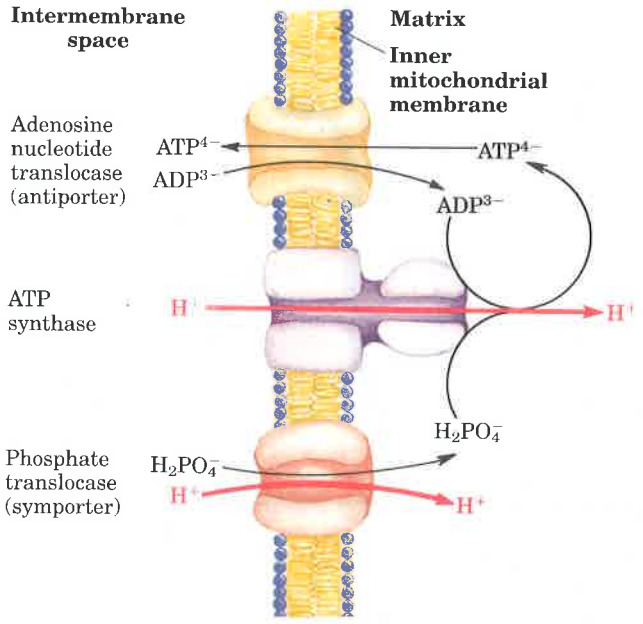



**Figure 18–24** Transport systems of the mitochondrial inner membrane carry ADP and $P_i$ into the matrix and allow the newly synthesized ATP to leave. The ATP–ADP translocase is an antiporter; the same protein moves ADP into the matrix and ATP out. The effect of replacing $ATP^{4-}$ with $ADP^{3-}$ is the net efflux of one negative charge, which is favored because the matrix is electrically negative relative to the outside. At pH 7, $P_i$ is present as both $HPO_4^{2-}$ and $H_2PO_4^-$; the transport system that carries $P_i$ into the matrix is specific for $H_2PO_4^-$. There is no net flow of charge during symport of $H_2PO_4^-$ and $H^+$, but the relatively low proton concentration in the matrix favors the inward movement of $H^+$. Thus the proton-motive force is responsible both for providing the energy for ATP synthesis by ATP synthase and for transporting substrates (ADP and $P_i$) in and product (ATP) out of the mitochondrial matrix.

Adenine nucleotide translocase is specifically inhibited by atractyloside, a toxic glycoside formed by a species of thistle; for centuries it has been known that grazing cattle are poisoned when they ingest this plant. If the transport of ADP into and ATP out of the mitochondria is inhibited, cytosolic ATP cannot be regenerated from ADP, explaining the toxicity of atractyloside (Table 18–4).

A second membrane transport system essential to oxidative phosphorylation is the **phosphate translocase,** which promotes symport of one $H_2PO_4^-$ and one $H^+$ into the matrix. This transport process, too, is favored by the transmembrane proton gradient (Fig. 18–24).

## Shuttle Systems Are Required for Mitochondrial Oxidation of Cytosolic NADH

The NADH dehydrogenase of the inner mitochondrial membrane of animal cells can accept electrons only from NADH in the matrix. Given that the inner membrane is not permeable to cytosolic NADH, how can the NADH generated by glycolysis outside mitochondria be reoxidized to $NAD^+$ by $O_2$ via the respiratory chain? Special shuttle systems carry

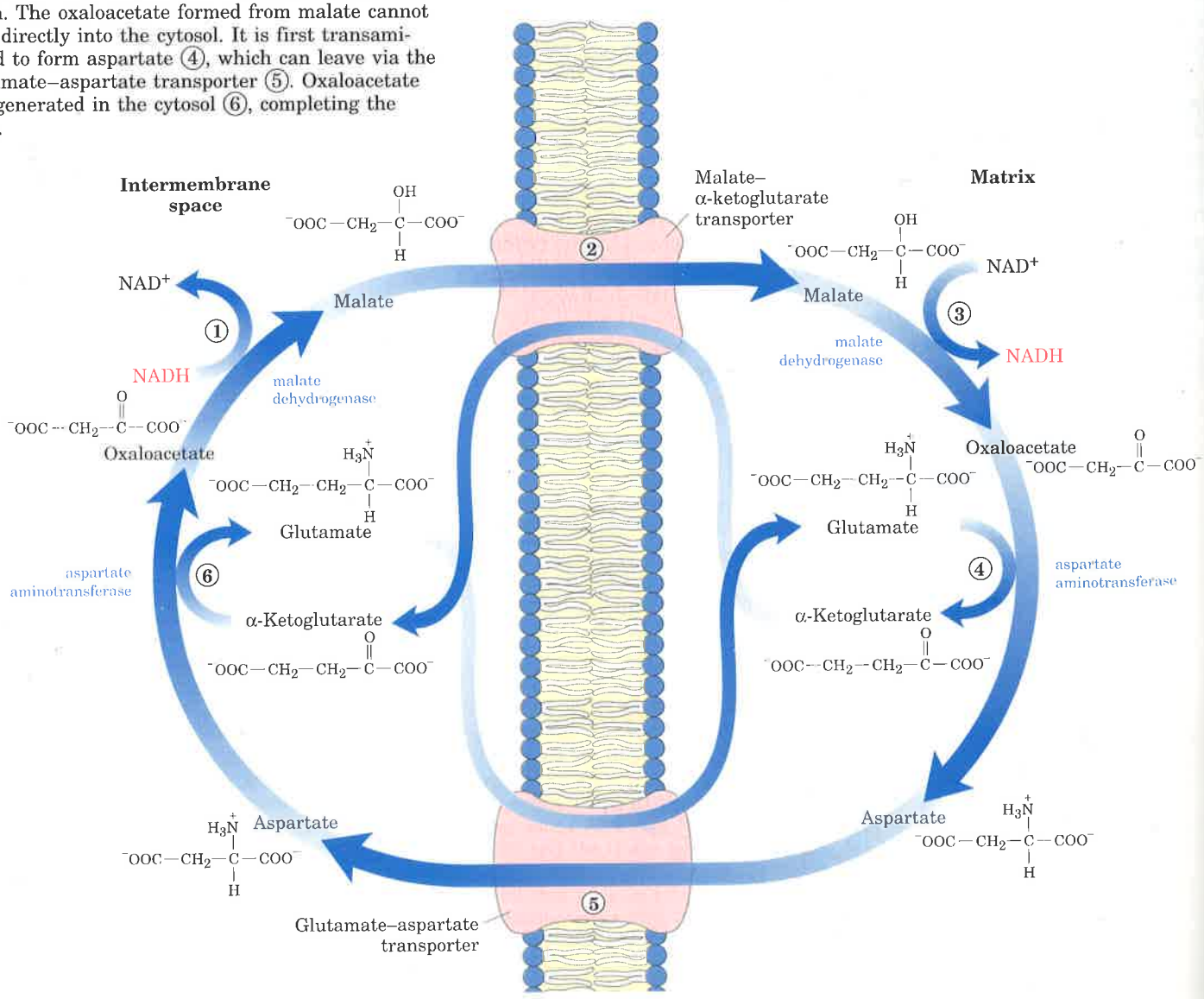



**Figure 18–25** The malate–aspartate shuttle for transporting reducing equivalents from cytosolic NADH into the mitochondrial matrix. ① NADH in the cytosol (intermembrane space) passes two reducing equivalants to oxaloacetate, producing malate. ② Malate is transported across the inner membrane by the malate–α-ketoglutarate transporter. ③ In the matrix, malate passes two reducing equivalants to $NAD^+$; the resulting matrix NADH is oxidized by the mitochondrial respiratory chain. The oxaloacetate formed from malate cannot pass directly into the cytosol. It is first transaminated to form aspartate ④, which can leave via the glutamate–aspartate transporter ⑤. Oxaloacetate is regenerated in the cytosol ⑥, completing the cycle.

reducing equivalents from cytosolic NADH into mitochondria by an indirect route. The most active NADH shuttle, which functions in liver, kidney, and heart mitochondria, is the **malate–aspartate shuttle** (Fig. 18–25). The reducing equivalents of cytosolic NADH are first transferred to cytosolic oxaloacetate to yield malate by the action of cytosolic malate dehydrogenase. The malate thus formed passes through the inner membrane into the matrix via the malate–$\alpha$-ketoglutarate transport system. Within the matrix the reducing equivalents are passed by the action of matrix malate dehydrogenase to matrix $NAD^+$, forming NADH; this NADH can then pass electrons directly to the respiratory chain in the inner membrane. Three molecules of ATP are generated as this pair of electrons passes to $O_2$. Cytosolic oxaloacetate must be regenerated via transamination reactions (see Fig. 17–5) and the activity of membrane transporters (Fig. 18–25) to start another cycle of the shuttle.

In skeletal muscle and brain, another type of NADH shuttle, the **glycerol-3-phosphate shuttle,** occurs (Fig. 18–26). It differs from the malate–aspartate shuttle in that it delivers the reducing equivalents from NADH into Complex III, not Complex I (Fig. 18–9), providing only enough energy to synthesize two ATP molecules per pair of electrons.

The mitochondria of higher plants have an externally oriented NADH dehydrogenase that is able to transfer electrons directly from cytosolic NADH into the respiratory chain.

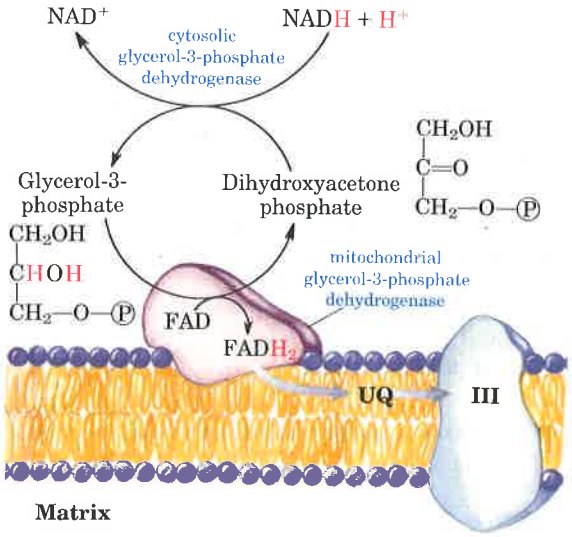



**Figure 18–26** The glycerol-3-phosphate shuttle, an alternative means of moving reducing equivalents from the cytosol to the mitochondrial matrix. Dihydroxyacetone phosphate in the cytosol accepts two reducing equivalents from cytosolic NADH in a reaction catalyzed by cytosolic glycerol-3-phosphate dehydrogenase. A membrane-bound isozyme of glycerol-3-phosphate dehydrogenase, located on the outer face of the inner membrane, transfers two reducing equivalents from glycerol-3-phosphate in the intermembrane space to ubiquinone. Note that this shuttle does not involve membrane transport systems.

## Oxidative Phosphorylation Produces Most of the ATP Made in Aerobic Cells

The complete oxidation of a molecule of glucose to $CO_2$ yields two ATP and two NADH from glycolysis in the cytosol (Chapter 14); two NADH from pyruvate oxidation in the mitochondrial matrix (Chapter 15); and two ATP, six NADH, and two $FADH_2$ from citric acid cycle reactions in the matrix (Chapter 15). Each NADH produced in the matrix yields three ATP from mitochondrial oxidative phosphorylation, and for each $FADH_2$, two ATP are generated. Cytosolic NADH, after shuttling into the matrix, yields two or three ATP, depending on which shuttle is used. The total yield of glucose oxidation is therefore 36 or 38 ATP per glucose (Table 18–5).

**Table 18–5 ATP yield from complete oxidation of glucose**

| Process | Direct product | Final ATP |
|---|---|---|
| Glycolysis | 2 NADH (cytosolic) | 4 or 6* |
| | 2 ATP | 2 |
| Pyruvate oxidation (two per glucose) | 2 NADH (mitochondrial matrix) | 6 |
| Acetyl-CoA oxidation (two per glucose) | 6 NADH (mitochondrial matrix) | 18 |
| | 2 $FADH_2$ | 4 |
| | 2 ATP or 2 GTP | 2 |
| Total yield per molecule of glucose | | 36 or 38 |

*The number depends on which shuttle system is used to transfer reducing equivalents into the mitochondrial matrix.

**Table 18–6 ATP yield from complete oxidation of palmitoyl-CoA***

| Process | Direct product | Final ATP |
|---|---|---|
| $\beta$ Oxidation | 7 NADH | 21 |
| | 7 $FADH_2$ | 14 |
| | 8 Acetyl-CoA | |
| Citric acid cycle | 24 NADH | 72 |
| | 8 $FADH_2$ | 16 |
| | 8 ATP or 8 GTP | 8 |
| Total yield per molecule of palmitoyl-CoA | | 131 |

* The activation of palmitate to palmitoyl-CoA costs two ATP.

By comparison, glycolysis under anaerobic conditions (lactate fermentation) yields only two ATP per glucose. Clearly, the evolution of oxidative phosphorylation provided a tremendous increase in the energetic efficiency of catabolism.

The oxidation of fatty acids and amino acids also takes place within the mitochondrial matrix, and the $FADH_2$ and NADH produced by these oxidative pathways also serve as electron donors for oxidative phosphorylation. Complete oxidation of the 16-carbon saturated fatty acid palmitate to $CO_2$ produces eight acetyl-CoAs, seven $FADH_2$, and seven NADH (Chapter 16). The oxidation of each acetyl-CoA via the citric acid cycle produces three NADH, one $FADH_2$, and one ATP (GTP). The gain in ATP is therefore 131 ATP (Table 18–6); but, because the activation of palmitate to palmitoyl-CoA costs two ATP equivalents, the net gain is 129 ATP per palmitate. A similar calculation may be made for the ATP yield upon oxidation of each of the amino acids (Chapter 17). The important conclusion here is that aerobic oxidative pathways that result in electron transfer to $O_2$ accompanied by oxidative phosphorylation account for the vast majority of the ATP produced in catabolism.

## Oxidative Phosphorylation Is Regulated by Cellular Energy Needs

The rate of respiration ($O_2$ consumption) in mitochondria is under tight regulation; it is generally limited by the availability of ADP as a substrate for phosphorylation. As we saw in Figure 18–13b, the respiration rate in isolated mitochondria is low in the absence of ADP and increases strikingly with the addition of ADP; this phenomenon is part of the definition of coupling of oxidation and phosphorylation. The dependence of the rate of $O_2$ consumption on the concentration of the $P_i$ acceptor ADP, called **acceptor control** of respiration, can be dramatic. In some animal tissues the **acceptor control ratio,** the ratio of the maximal rate of ADP-induced $O_2$ consumption to the basal rate in the absence of ADP, is at least 10.

The intracellular concentration of ADP is one measure of the energy status of cells. Another, related measure is the **mass-action ratio** of the ATP–ADP system: $[ATP]/([ADP][P_i])$. Normally this ratio is very high, so that the ATP–ADP system is almost fully phosphorylated. When the rate of some energy-requiring process in cells (protein synthesis, for example) increases, there is an increased rate of breakdown of ATP to ADP and $P_i$, lowering the mass-action ratio. With more ADP available for oxidative phosphorylation, the rate of respiration increases, causing regeneration of ATP. This continues until the mass-action ratio returns to its normal high level, at which point respiration slows again. The rate of oxidation of cell fuels is regulated with such sensitivity and precision that the ratio $[ATP]/([ADP][P_i])$ fluctuates only slightly in most tissues, even during extreme variations in energy demand. In short, ATP is formed only as fast as it is used in energy-requiring cell activities.

## Uncoupled Mitochondria in Brown Fat Produce Heat

There is a remarkable and instructive exception to the general rule that respiration slows when the ATP supply is adequate. In most mammals, including humans, newborns have a type of adipose tissue called

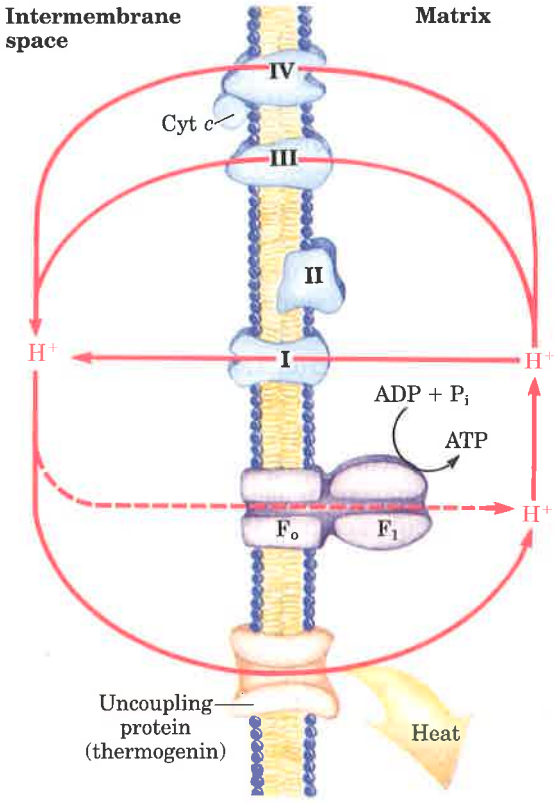



**Figure 18–27** The uncoupling protein (thermogenin) of brown fat mitochondria, by providing an alternative route for protons to reenter the mitochondrial matrix, causes the energy conserved by proton pumping to be dissipated as heat.

**brown fat** in which fuel oxidation serves not to produce ATP, but to generate heat to keep the newborn warm. Hibernating mammals (grizzly bears, for example) also have large amounts of brown fat. This specialized adipose tissue, located at the back of the neck of human infants, is brown because of the presence of large numbers of mitochondria and thus large amounts of cytochromes, whose heme groups are strong absorbers of visible light.

The mitochondria of brown fat oxidize fuels (particularly fatty acids) normally, passing electrons through the respiratory chain to $O_2$. This electron transfer is accompanied by proton pumping out of the matrix, as in other mitochondria. The mitochondria of brown fat, however, have a unique protein in their inner membrane: **thermogenin,** also called the **uncoupling protein** (UCP) (Table 18–4). This protein, an integral membrane protein, provides a path for protons to return to the matrix without passing through the $F_oF_1$ complex (Fig. 18–27). As a result of this short-circuiting of protons, the energy of oxidation is not conserved by ATP formation but is dissipated as heat, which contributes to maintaining the body temperature. For the hairless newborn infant, maintaining body heat is an important use of metabolic energy. Hibernating animals depend on uncoupled mitochondria of brown fat to generate heat during their long winter period of dormancy (see Box 16–1).

## ATP-Producing Pathways Are Coordinately Regulated

The major catabolic pathways (glycolysis, the citric acid cycle, fatty acid and amino acid oxidation, and oxidative phosphorylation) have interlocking and concerted regulatory mechanisms that allow them to

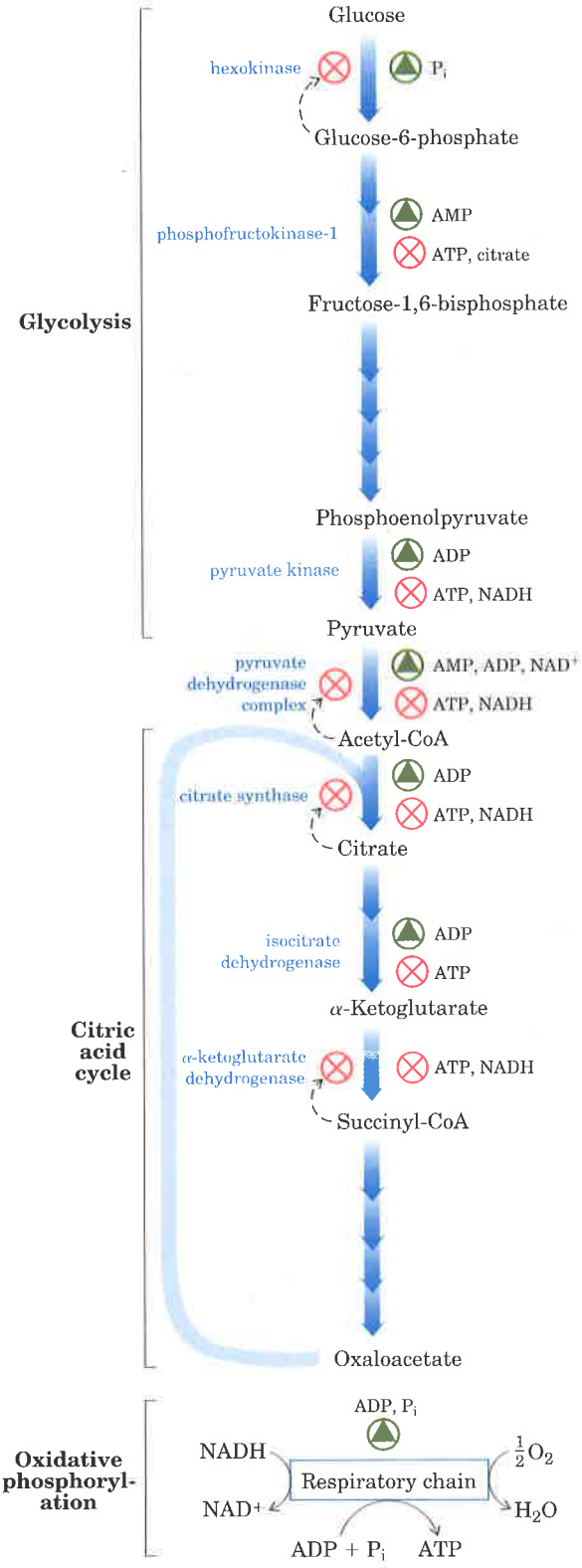



**Figure 18–28** Interlocking regulation of glycolysis, pyruvate oxidation, the citric acid cycle, and oxidative phosphorylation by the relative concentrations of ATP, ADP, and AMP, and by NADH. High [ATP] (or low [ADP] and [AMP]) produces low rates of glycolysis, pyruvate oxidation, acetate oxidation via the citric acid cycle, and oxidative phosphorylation. All four pathways are accelerated when there is an increase in the rate of ATP utilization and increased formation of ADP, AMP, and $P_i$. Interlocking of glycolysis and the citric acid cycle by citrate, which inhibits glycolysis, supplements the action of the adenine nucleotide system. In addition, increased levels of NADH and acetyl-CoA also inhibit the oxidation of pyruvate to acetyl-CoA, and high [NADH]/[$NAD^+$] ratios inhibit the dehydrogenase reactions of the citric acid cycle (p. 469).

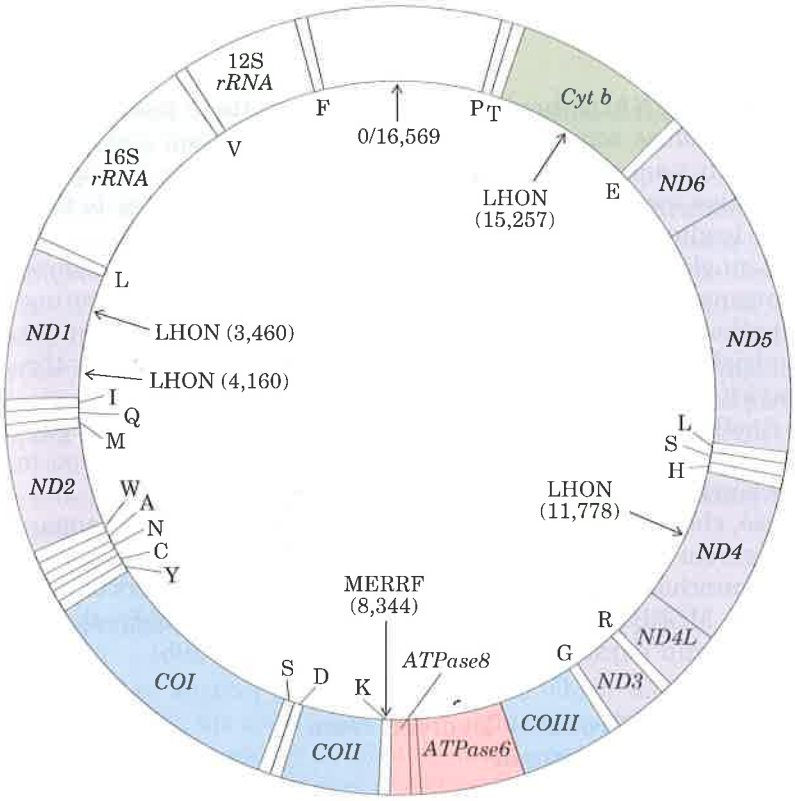
function together in an economical and self-regulating manner to produce ATP and biosynthetic precursors. The relative concentrations of ATP and ADP control not only the rates of electron transfer and oxidative phosphorylation but also the rates of the citric acid cycle, pyruvate oxidation, and glycolysis (Fig. 18–28). Whenever there is an increased drain on ATP, the rate of electron transfer and oxidative phosphorylation increases. Simultaneously, the rate of pyruvate oxidation via the citric acid cycle increases, thus increasing the flow of electrons into the respiratory chain. These events can in turn evoke an increase in the rate of glycolysis, increasing the rate of pyruvate formation. When the conversion of ADP to ATP lowers the ADP concentration, acceptor control will slow electron transfer and thus oxidative phosphorylation. Glycolysis and the citric acid cycle will also slow, because ATP is an allosteric inhibitor of phosphofructokinase-1 (see Fig. 14–20) and of pyruvate dehydrogenase (see Fig. 15–14).

Phosphofructokinase-1 is inhibited not only by ATP but by citrate, the first intermediate of the citric acid cycle. When both ATP and citrate are elevated they produce a concerted allosteric inhibition that is greater than the sum of their individual effects.

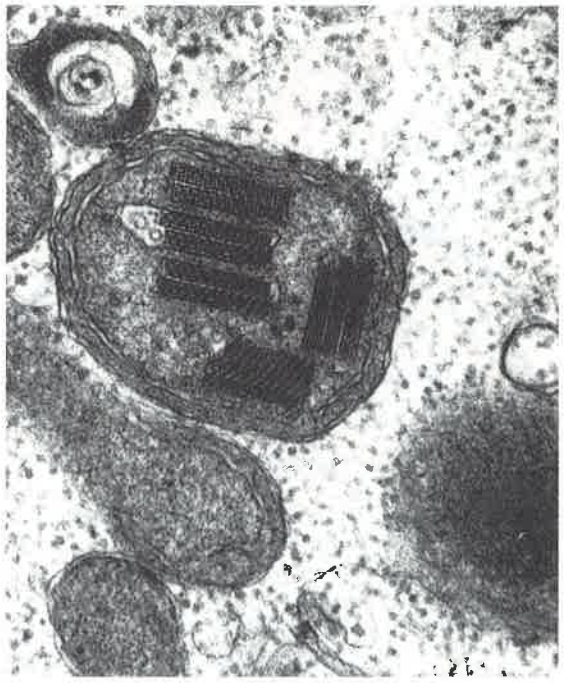
In many types of tumor cells, this interlocking coordination appears to be defective: glycolysis proceeds at a higher rate than required by the citric acid cycle. As a result, cancer cells use far more blood glucose than normal cells, but cannot oxidize the excess pyruvate formed by rapid glycolysis, even in the presence of $O_2$. To reoxidize cytoplasmic NADH, most of the pyruvate is reduced to lactate (Chapter 14), which passes from the cells into the blood. The high glycolytic rate may result in part from smaller numbers of mitochondria in tumor cells. In addition, some tumor cells overproduce an isozyme of hexokinase that associates with the cytosolic face of the mitochondrial inner membrane and is insensitive to feedback inhibition by glucose-6-phosphate (p. 432). This enzyme may monopolize the ATP produced in mitochondria, using it to produce glucose-6-phosphate and committing the cell to continue glycolysis.

## Mutations in Mitochondrial Genes Cause Human Disease

Mitochondria contain their own genome, a circular, double-stranded DNA molecule. There are 37 genes (16,569 base pairs) in the human mitochondrial chromosome (Fig. 18–29), including 13 that encode proteins of the respiratory chain (Table 18–7); the remaining genes code for ribosomal and transfer RNA molecules essential to the protein-synthesizing machinery of mitochondria. Many of the mitochondrial proteins are encoded in nuclear genes, synthesized on cytoplasmic ribosomes, then imported and assembled within mitochondria.




(a)



(b)

**Figure 18–29** **(a)** Map of human mitochondrial DNA, showing the genes that encode proteins of Complex I, the NADH dehydrogenase (*ND1* to *ND6*); the cytochrome *b* of Complex III (*Cyt b*); the subunits of cytochrome oxidase (Complex IV) (*COI* to *COIII*); and two subunits of ATP synthase (*ATPase6* and *ATPase8*). The colors of the genes correspond to the colors of the complexes shown in Fig. 18–8. Also shown are the genes for ribosomal RNAs (*rRNA*) and for a number of mitochondrion-specific transfer RNAs. Transfer RNA specificity is indicated by the one-letter codes for amino acids. Arrows indicate the positions where alterations of base sequence (mutations) are known to cause Leber's hereditary optic neuropathy (LHON) and myoclonic epilepsy and ragged-red fiber disease (MERRF). The numbers in parentheses indicate the position of the altered nucleotides; nucleotide number 1 is at the top of the circle. **(b)** Electron micrograph of an abnormal mitochondrion from the muscle of an individual with MERRF, showing the paracrystalline protein inclusions sometimes present in the mutant mitochondria.

**Table 18–7 Respiratory proteins encoded in the human mitochondrial chromosome**

| Complex | Total number of subunits | Number of subunits encoded in mitochondrial DNA |
|---|---|---|
| I NADH dehydrogenase | >25 | 7 |
| II Succinate dehydrogenase | 4 | 0 |
| III Ubiquinone-cytochrome *c* oxidoreductase | 9 | 1 |
| IV Cytochrome oxidase | 13 | 3 |
| V ATP synthase | 12 | 2 |

Mutations in the mitochondrial DNA are known to cause the human disease called **Leber's hereditary optic neuropathy** (LHON). This rare genetic disease affects the central nervous system, including the optic nerves, causing bilateral loss of vision, of rapid onset, in early adulthood. The disease is invariably inherited from the female parent, a consequence of the fact that all of the mitochondria of the developing embryo are derived from the mother. The unfertilized egg contains many mitochondria, and the few mitochondria in the much smaller sperm cell do not enter the egg at the time of fertilization. Mitochondria arise only from the division of preexisting mitochondria.

A single base change in the mitochondrial gene *ND4* (Fig. 18–29a) changes an Arg residue to a His residue in one of the proteins of Complex I, and the result is mitochondria partially defective in electron

transfer from NADH to ubiquinone. Although these mitochondria are able to produce some ATP by electron transfer from succinate, they apparently cannot supply ATP in sufficient quantity to support the very active metabolism of neurons. One result is damage to the optic nerve, leading to blindness.

A single base change in the mitochondrial gene for cytochrome *b*, a component of Complex III, also produces LHON, demonstrating that the pathology of the disease results from a general reduction of mitochondrial function, not just from a defect in electron transfer through Complex I.

Another serious human genetic disease, **myoclonic epilepsy and ragged-red fiber disease** (MERRF) is caused by a mutation in the mitochondrial gene that encodes a transfer RNA (Fig. 18–29). This disease, characterized by uncontrollable muscular jerking, apparently results from defective production of several of the proteins synthesized using mitochondrial transfer RNAs. Skeletal muscle fibers of individuals with MERRF have abnormally shaped mitochondria that sometimes contain paracrystalline structures (Fig. 18–29b).

## Mitochondria Probably Evolved from Endosymbiotic Bacteria

The fact that mitochondria contain their own DNA, ribosomes, and transfer RNAs supports the theory of the endosymbiotic origin of mitochondria (see Fig. 2–17). This theory supposes that the first organisms capable of aerobic metabolism, including respiration-linked ATP production, were prokaryotes. Primitive eukaryotes that lived anaerobically (by fermentation) acquired the ability to carry out oxidative phosphorylation when they established a symbiotic relationship with bacteria living in their cytosol. After much evolution and the movement of many bacterial genes into the nucleus of the "host" eukaryote, the endosymbiotic bacteria eventually became mitochondria.

This theory presumes that early free-living prokaryotes had the enzymatic machinery for oxidative phosphorylation, and it predicts that their modern prokaryotic descendents have respiratory chains closely similar to those of modern eukaryotes. They do.

Aerobic bacteria carry out NAD-linked electron transfer from substrates to $O_2$, coupled to the phosphorylation of cytosolic ADP. The dehydrogenases are located in the bacterial cytosol, but the electron carriers of the respiratory chain are in the plasma membrane. The electron carriers are similar to those of mitochondria, act in the same sequence (Fig. 18–30), and translocate protons outward across the plasma membrane concomitantly with electron transfer to $O_2$. We noted earlier that bacteria such as *E. coli* have $F_oF_1$ complexes in their plasma membranes; the $F_1$ portions protrude into the cytosol and cata-

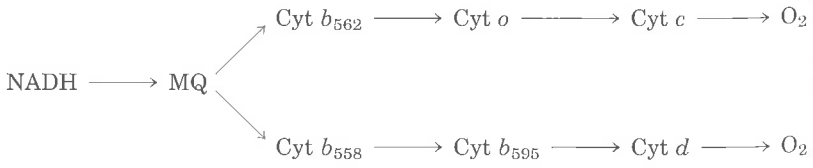



**Figure 18–30** Respiratory chain in the inner membrane of *E. coli*. Electrons from NADH pass to menaquinone (MQ), at which point the chain branches. The upper path is dominant in cells grown under normal aerobic conditions, but when $O_2$ is the limiting factor in growth, the lower route predominates.

lyze ATP synthesis from ADP and $P_i$ as protons flow back into the cell through proton channels formed by $F_o$.

Certain bacterial transport systems bring about uptake of extracellular nutrients (lactose, for example) against a concentration gradient, in symport with protons (see Fig. 10–25). The respiration-linked transmembrane proton extrusion provides the driving force for this uptake. The rotary motion of bacterial flagella, which move cells through their surroundings, is provided by "proton turbines," molecular rotary motors driven not by ATP but directly by the transmembrane electrochemical potential generated by respiration-linked proton pumping (Fig. 18–31). It appears likely that the chemiosmotic mechanism evolved early (before the emergence of eukaryotes). The proton-motive force can clearly be used to power processes other than ATP synthesis.

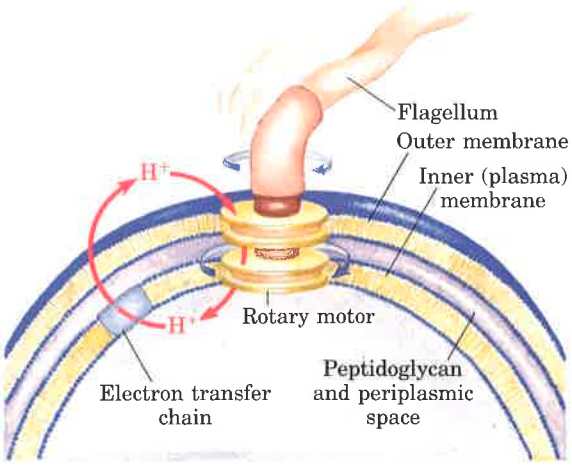



**Figure 18–31** Rotation of bacterial flagella by proton-motive force. The shaft and rings at the base of the flagellum make up a rotary motor that has been called a "proton turbine." Protons ejected by electron transfer flow back into the cell through the "turbine," causing rotation of the shaft of the flagellum. This motion differs fundamentally from the motion of muscle or of eukaryotic flagella and cilia, for which ATP hydrolysis is the energy source.

## Photosynthesis: Harvesting Light Energy

We now turn to another reaction sequence in which the flow of electrons is coupled to the synthesis of ATP: light-driven phosphorylation. The capture of solar energy by photosynthetic organisms and its conversion into the chemical energy of reduced organic compounds is the ultimate source of nearly all biological energy. Photosynthetic and heterotrophic organisms live in a balanced steady state in the biosphere (Fig. 18–32). Photosynthetic organisms trap solar energy and form ATP and NADPH, which they use as energy sources to make carbohydrates and other organic components from $CO_2$ and $H_2O$; simultaneously, they release $O_2$ into the atmosphere. Aerobic heterotrophs (we humans, for example) use the $O_2$ so formed to degrade the energy-rich organic products of photosynthesis to $CO_2$ and $H_2O$, generating ATP for their own activities. The $CO_2$ formed by respiration in heterotrophs returns to the atmosphere, to be used again by photosynthetic organisms. Solar energy thus provides the driving force for the continuous cycling of atmospheric $CO_2$ and $O_2$ through the biosphere and provides the reduced substrates (fuels), such as glucose, on which nonphotosynthetic organisms depend.

Enormous amounts of energy are stored as products of photosynthesis. Each year at least $10^{17}$ kJ of free energy from sunlight is captured and used for biosynthesis by photosynthetic organisms. This is more than ten times the fossil-fuel energy used each year by people the world over. Even fossil fuels (coal, oil, and natural gas) are the products of photosynthesis that took place millions of years ago. Because of our global dependence upon solar energy, past and present, for both energy and food, discovering the mechanism of photosynthesis is a central goal of biochemical research.

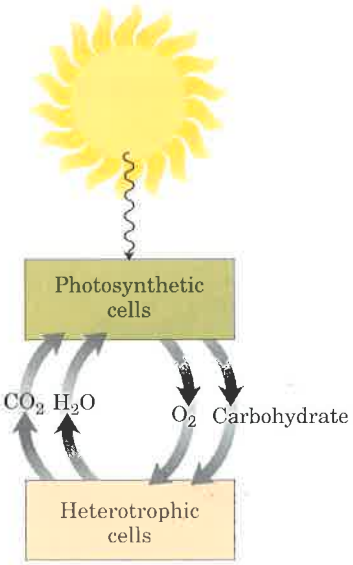



**Figure 18–32** Solar energy is the ultimate source of all biological energy. Photosynthetic organisms use the energy of sunlight to manufacture glucose and other organic cell products, which heterotrophic cells use as energy and carbon sources.

The overall equation of photosynthesis describes an oxidation–reduction reaction in which $H_2O$ donates electrons (as hydrogen) for the reduction of $CO_2$ to carbohydrate ($CH_2O$):

$$CO_2 + H_2O \xrightarrow{\text{light}} O_2 + (CH_2O)$$

Unlike NADH (the hydrogen donor in oxidative phosphorylation), $H_2O$ is a poor electron donor; its standard reduction potential is 0.82 V, compared with −0.32 V for NADH. The central difference between photophosphorylation and oxidative phosphorylation is that the latter process begins with a good electron donor, whereas the former requires the input of energy in the form of light to *create* a good electron donor. Except for this crucial difference, the two processes are remarkably similar. In photophosphorylation, electrons flow through a series of membrane-bound carriers including cytochromes, quinones, and iron–sulfur proteins, while protons are pumped across a membrane to create an electrochemical potential. This potential is the driving force for ATP synthesis from ADP and $P_i$ by a membrane-bound ATP synthase complex closely similar to that which functions in oxidative phosphorylation.

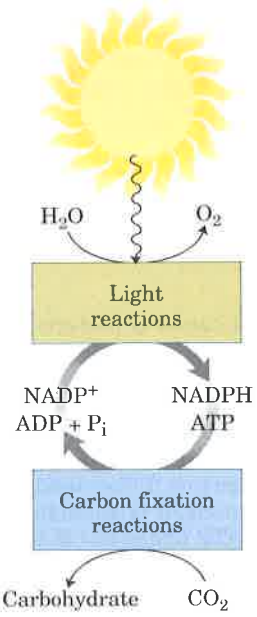


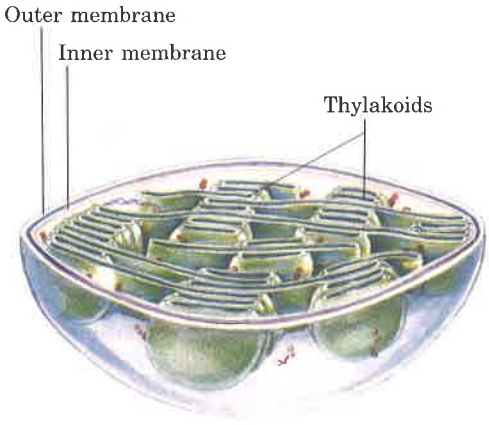

**Figure 18–33** The light reactions generate energy-rich NADPH and ATP at the expense of solar energy. These products are used in the carbon fixation reactions, which occur in light or darkness, to reduce $CO_2$ to form trioses and more complex compounds (such as glucose) derived from trioses.

Photosynthesis encompasses two processes: the **light reactions,** which occur only when plants are illuminated, and the **carbon fixation reactions,** or so-called dark reactions, which occur in both light and darkness (Fig. 18–33). In the light reactions, chlorophyll and other pigments of the photosynthetic cells absorb light energy and conserve it in chemical form as the two energy-rich products ATP and NADPH; simultaneously, $O_2$ is evolved. In the carbon fixation reactions, ATP and NADPH are used to reduce $CO_2$ to form glucose and other organic products. The formation of $O_2$, which occurs only in the light, and the reduction of $CO_2$, which does not require light, thus are distinct and separate processes. In this chapter we are concerned only with the light reactions; the reduction of $CO_2$ is described in Chapter 19.

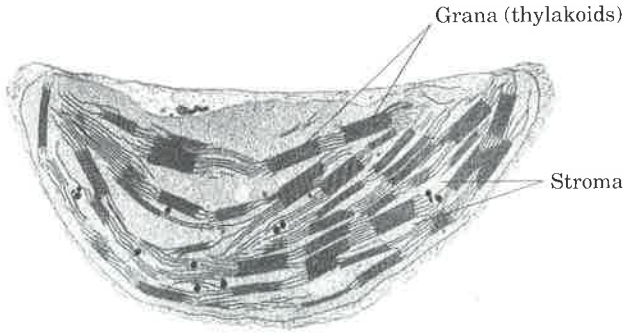
In photosynthetic eukaryotic cells, both the light and carbon fixation reactions take place in the chloroplasts (Fig. 18–34). When solar energy is not available, mitochondria in the plant cell generate ATP by oxidizing carbohydrates originally produced in chloroplasts in the light.

Chloroplasts may assume many different shapes in different species, and they usually have a much larger volume than mitochondria.

**Figure 18–34** Schematic diagram **(a)** and electron micrograph **(b)** of a single chloroplast at high magnification.




(a)




(b)

They are surrounded by a continuous outer membrane, which, like the outer mitochondrial membrane, is permeable to small molecules and ions. An inner membrane system encloses the internal compartment, in which there are many flattened, membrane-surrounded vesicles or sacs, called **thylakoids,** which are usually arranged in stacks called **grana.** The thylakoid membranes are separate from the inner chloroplast membrane. Embedded in the thylakoid membranes are the photosynthetic pigments and all the enzymes required for the primary light reactions. The fluid in the compartment surrounding the thylakoids, the **stroma,** contains most of the enzymes required for the carbon fixation reactions, in which $CO_2$ is reduced to form triose phosphates and, from them, glucose.

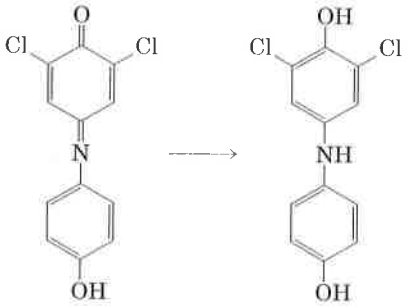
Our discussion here focuses on the nature of the light-absorbing systems and their roles in photosynthesis.

## Light Produces Electron Flow in Chloroplasts

How do the pigment molecules of the thylakoid membranes transduce absorbed light energy into chemical energy? The key to answering this question came from a discovery made in 1937 by Robert Hill, a pioneer in photosynthesis research. He found that when leaf extracts containing chloroplasts were supplemented with a nonbiological hydrogen acceptor and then illuminated, evolution of $O_2$ and simultaneous reduction of the hydrogen acceptor took place, according to an equation now known as the **Hill reaction:**

$$2H_2O + 2A \xrightarrow{\text{light}} 2AH_2 + O_2$$

where A is the artificial hydrogen acceptor. One of the nonbiological hydrogen acceptors used by Hill was the dye 2,6-dichlorophenolindophenol, now called a **Hill reagent,** which in its oxidized form (A) is blue and in its reduced form ($AH_2$) is colorless. When the leaf extract supplemented with the dye was illuminated, the blue dye became colorless and $O_2$ was evolved. In the dark neither $O_2$ evolution nor dye reduction took place. This was the first specific clue to how absorbed light energy is converted into chemical energy: it causes electrons to flow from $H_2O$ to an electron acceptor. Moreover, Hill found that $CO_2$ was not required for this reaction, nor was it reduced to a stable form under these conditions. He therefore concluded that $O_2$ evolution can be dissociated from $CO_2$ reduction. Several years later it was found that $NADP^+$ is the biological electron acceptor in chloroplasts, according to the equation



Oxidized form (blue) — Reduced form (colorless)

Dichlorophenolindophenol

$$2H_2O + 2NADP^+ \xrightarrow{\text{light}} 2NADPH + 2H^+ + O_2$$

This equation shows an important distinction between mitochondrial oxidative phosphorylation and the analogous process in chloroplasts: in chloroplasts electrons flow from $H_2O$ to $NADP^+$, whereas in mitochondrial respiration electrons flow in the opposite direction, from NADH or NADPH to $O_2$, with the release of free energy. Because light-induced electron flow in chloroplasts is in the reverse or "uphill" direction, from $H_2O$ to $NADP^+$, it cannot occur without the input of free energy; this energy comes from light. To understand how this occurs, we must first consider the effects of light absorption on molecular structure.

## Absorption of Light Excites Molecules

Visible light is electromagnetic radiation of wavelengths 400 to 700 nm, a small part of the electromagnetic spectrum (Fig. 18–35), ranging from violet to red. The energy of a single **photon** (a quantum of light) is greater at the violet end of the spectrum than at the red end. The energy in a "mole" of photons (one einstein; $6 \times 10^{23}$ photons) is 170 to 300 kJ, about an order of magnitude more than the energy required to synthesize a mole of ATP from ADP and $P_i$ (about 30 kJ under standard conditions). The energy of an einstein in the infrared or microwave regions of the spectrum is too small to be useful in the kinds of photochemical events that occur in photosynthesis. UV light and x rays, on the other hand, have so much energy that they damage proteins and nucleic acids and induce mutations that are often lethal.

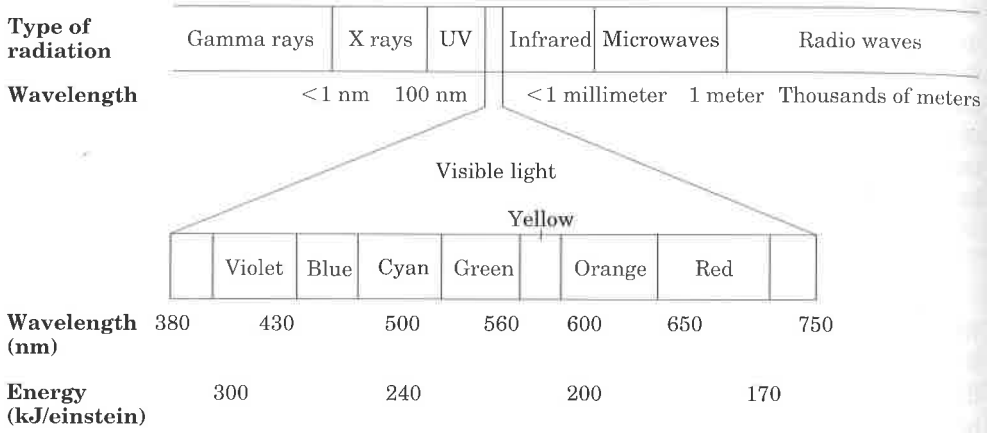
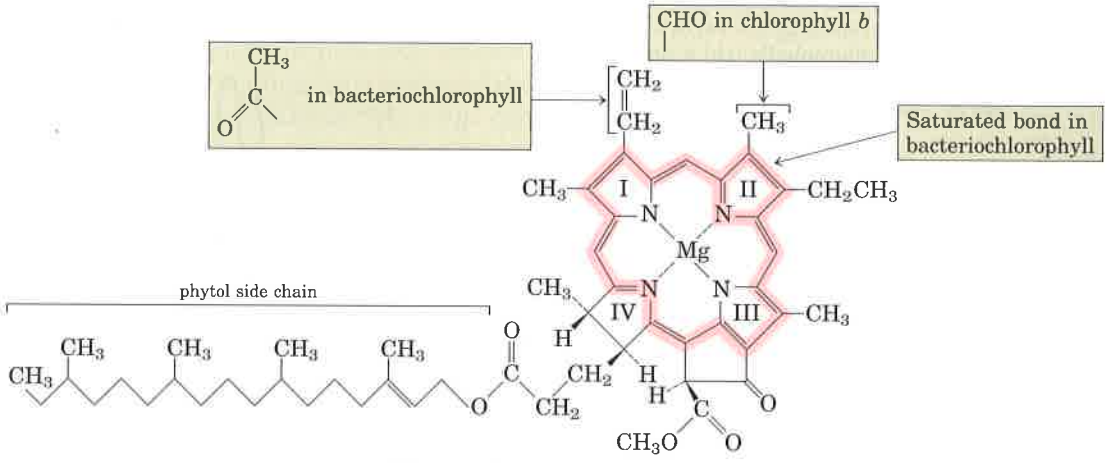



**Figure 18–35** The spectrum of electromagnetic radiation, and the energy of photons in the visible range of the spectrum. One einstein is $6 \times 10^{23}$ photons.

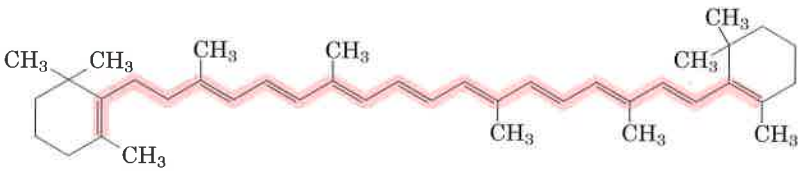
The ability of a molecule to absorb light depends upon the arrangement of electrons around the atomic nuclei in its structure. When a photon is absorbed, an electron is lifted to a higher energy level. This happens on an all-or-none basis; to be absorbed, the photon must contain a quantity of energy (a **quantum**) that exactly matches the energy of the electronic transition. A molecule that has absorbed a photon is in an **excited state,** which is generally unstable. The electrons lifted into higher-energy orbitals usually return rapidly to their normal lower-energy orbitals; the excited molecule reverts (decays) to the stable **ground state,** giving up the absorbed quantum as light or heat or using it to do chemical work. The light emitted upon decay of excited molecules, called **fluorescence,** is always of a longer wavelength (lower energy) than the absorbed light. Excitation of molecules by light and their fluorescent decay are extremely fast processes, occurring in about $10^{-15}$ and $10^{-12}$ s, respectively. The initial events in photosynthesis are therefore very rapid.

## Chlorophylls Absorb Light Energy for Photosynthesis

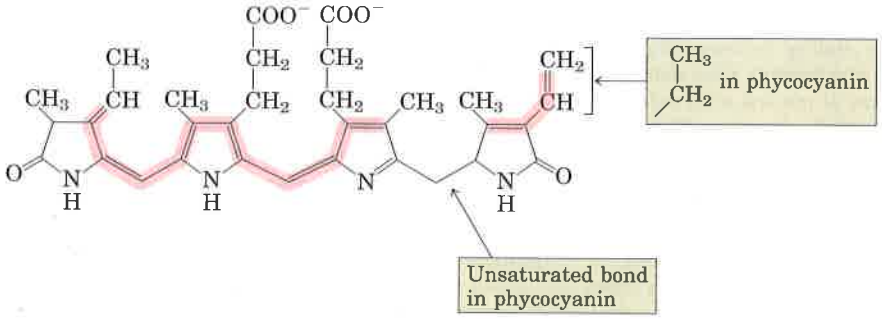
The most important light-absorbing pigments in the thylakoid membranes are the **chlorophylls,** green pigments with polycyclic, planar structures resembling the protoporphyrin of hemoglobin (see Fig. 7–18), except that $Mg^{2+}$, not $Fe^{2+}$, occupies the central position (Fig. 18–36). Chlorophyll *a*, present in the chloroplasts of all green plant cells, contains four substituted pyrrole rings, one of which (ring IV) is reduced, and a fifth ring that is not a pyrrole. All chlorophylls have a long **phytol** side chain, esterified to a carboxyl-group substituent in ring IV. The four inward-oriented nitrogen atoms of chlorophyll *a* are coordinated with the $Mg^{2+}$.




Chlorophyll *a*



β-Carotene




Phycoerythrin

**Figure 18–36** Structures of the primary photopigments chlorophylls *a* and *b* and bacteriochlorophyll, and of the accessory pigments β-carotene (a carotenoid) and phycoerythrin and phycocyanin (phycobilins). The areas shaded pink represent conjugated systems (alternating single and double bonds), which largely account for the absorption of visible light.

The heterocyclic five-ring system that surrounds the $Mg^{2+}$ has an extended polyene structure, with alternating single and double bonds. Such polyenes characteristically show strong absorption in the visible region of the spectrum (Fig. 18–37); the chlorophylls have unusually high molar absorption coefficients (see Box 5–1) and are therefore particularly well-suited for absorbing visible light during photosynthesis.

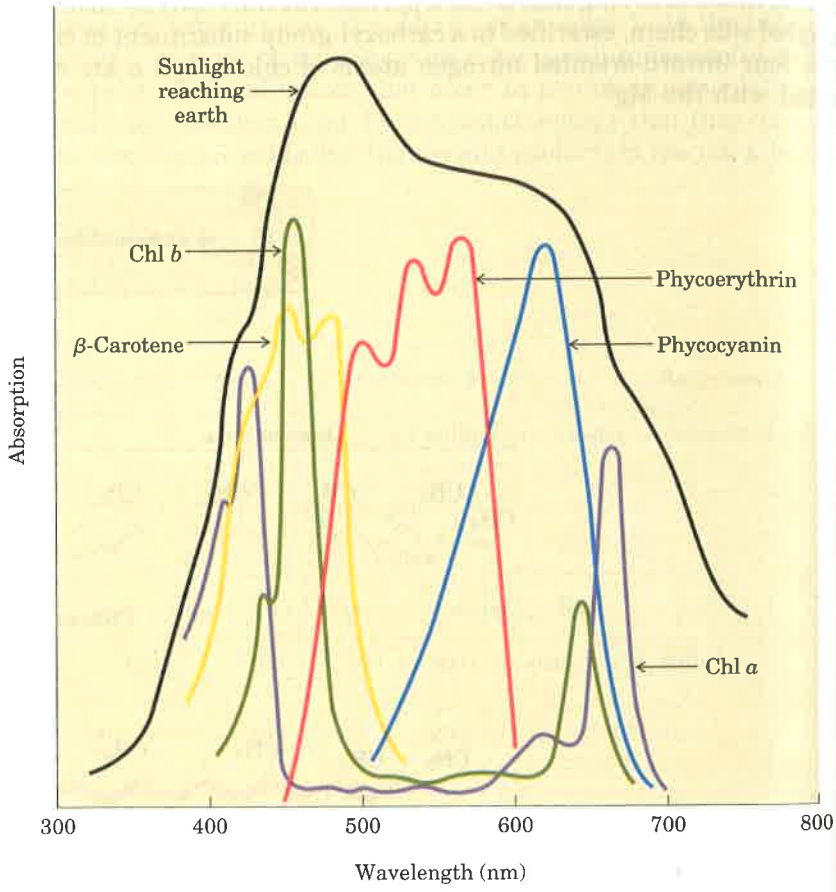



**Figure 18–37** Absorption of visible light by photopigments shown in Fig. 18–36. Plants are green because their pigments absorb light from the red and violet regions of the spectrum, leaving primarily green light to be reflected or transmitted. Compare the absorption spectra of the pigments with the spectrum of sunlight reaching the earth's surface; the combination of chlorophylls (chl *a* and chl *b*) and accessory pigments enables plants to harvest most of the energy available from sunlight.

Chloroplasts of higher plants always contain two types of chlorophyll. One is invariably chlorophyll *a*, and the second in many species is chlorophyll *b*, which has an aldehyde group instead of a methyl group attached to ring II (Fig. 18–36). Although both are green, their absorption spectra are slightly different (Fig. 18–37), allowing the two pigments to complement each other's range of light absorption in the visible region. Most higher plants contain about twice as much chlorophyll *a* as chlorophyll *b*. The bacterial chlorophylls differ only slightly from the plant pigments (Fig. 18–36).

### Accessory Pigments Also Absorb Light

In addition to chlorophylls, the thylakoid membranes contain secondary light-absorbing pigments, together called the **accessory pigments,** the carotenoids and phycobilins. **Carotenoids** may be yellow, red, or purple. The most important are **β-carotene** (Fig. 18–36), a red-orange isoprenoid compound that is the precursor of vitamin A in animals, and the yellow carotenoid **xanthophyll.** The carotenoid pig-

ments absorb light at wavelengths other than those absorbed by the chlorophylls (Fig. 18–37) and thus are supplementary light receptors. **Phycobilins** are linear tetrapyrroles that have the extended polyene system found in chlorophylls, but not their cyclic structure or central $Mg^{2+}$. Examples are phycoerythrin and phycocyanin (Fig. 18–36).

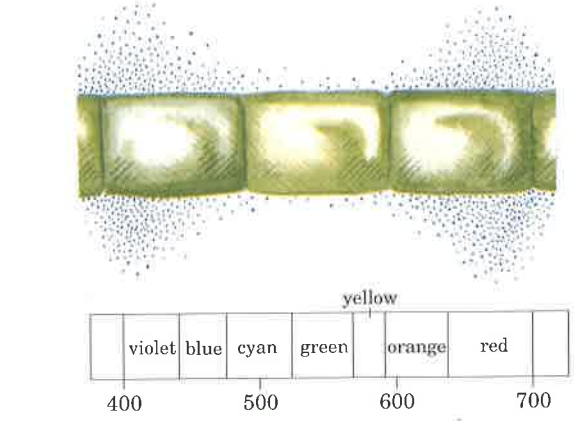
The relative amounts of the chlorophylls and the accessory pigments are characteristic for different plant species. It is variation in the proportions of these pigments that is responsible for the range of colors of photosynthetic organisms, which vary from the deep blue-green of spruce needles, to the greener green of maple leaves, to the red, brown, or even purple color of different species of multicellular algae and the leaves of some decorative plants.

Experimental determination of the effectiveness of light of different colors in promoting photosynthesis yields an **action spectrum** (Fig. 18–38), often useful in identifying the pigment primarily responsible for a biological effect of light. By capturing light in a region of the spectrum not used by other plants, a photosynthetic organism can claim its unique ecological niche. For example, the phycobilins, present only in red algae and cyanobacteria, absorb in the region 520 to 630 nm, allowing these organisms to live in niches where light of lower or higher wavelength has been filtered out by the pigments of other organisms living in the water above them, or by the water itself.

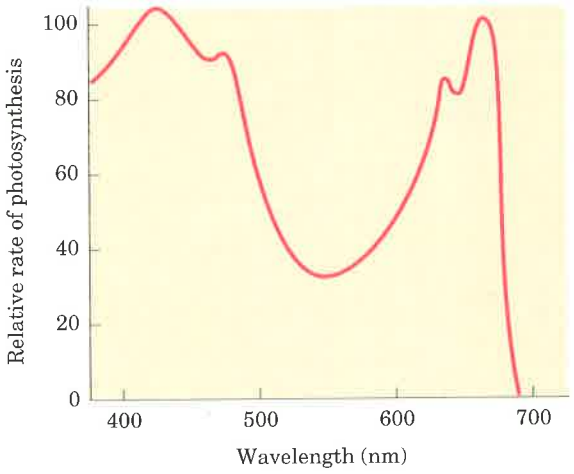
## Chlorophyll Funnels Absorbed Energy to Reaction Centers

The light-absorbing pigments of thylakoid membranes are arranged in functional sets or arrays called **photosystems.** In spinach chloroplasts each photosystem contains about 200 molecules of chlorophylls and about 50 molecules of carotenoids. The clusters can absorb light over the entire visible spectrum but especially well between 400 to 500 nm and 600 to 700 nm (Fig. 18–37). All the pigment molecules in a photosystem can absorb photons, but only a few can transduce the light energy into chemical energy. A transducing pigment consists of several chlorophyll molecules combined with a protein complex also containing tightly bound quinones; this complex is called a **photochemical reaction center.** The other pigment molecules in a photosystem are called **light-harvesting** or **antenna molecules.** They function to absorb light energy and transmit it at a very high rate to the reaction center where the photochemical reactions occur (Fig. 18–39, p. 578), which will be described in detail later.

The chlorophyll molecules in thylakoid membranes are bound to integral membrane proteins (chlorophyll *a*/*b*-binding, or CAB, proteins) that orient the chlorophyll relative to the plane of the membrane and confer light absorption properties that are subtly different from those of free chlorophyll. When isolated chlorophyll molecules in vitro are excited by light, the absorbed energy is quickly released as fluorescence and heat, but when chlorophyll in intact spinach leaves is excited by visible light (Fig. 18–40 (step ①), p. 579), very little fluorescence is observed. Instead, a direct transfer of energy from the excited chlorophyll (an antenna chlorophyll) to a neighboring chlorophyll molecule occurs, exciting the second molecule and allowing the first to return to its ground state (step ②). This **resonance energy transfer** is repeated to a third, fourth, or subsequent neighbor, until the chlorophyll at the photochemical reaction center becomes excited (step ③). In this special chlorophyll molecule, an electron is promoted by excitation to a higher-energy orbital. This electron then passes to a nearby electron




**(a)**




**(b)**

**Figure 18–38** Two ways to determine the action spectrum for photosynthesis. **(a)** The results of a classic experiment done by T.W. Englemann in 1882 to determine what wavelength of light was most effective in supporting photosynthesis. Englemann placed a filamentous, photosynthetic alga on a microscope stage and illuminated it with light from a prism, so that cells in one part of the filament received mainly blue light, another yellow, another red. To determine which cells carried out photosynthesis most actively, bacteria known to migrate toward regions of high $O_2$ concentration were also placed on the microscope slide. The distribution of bacteria showed highest $O_2$ levels (produced by photosynthesis) in the regions illuminated with violet and red light. **(b)** A similar experiment using modern techniques for the measurement of $O_2$ production yields the same result. An action spectrum describes the relative rate of photosynthesis for illumination with a constant number of photons of different wavelengths. Such an action spectrum is useful because it suggests (by comparison with absorption spectra such as those in Fig. 18–37) which pigments are able to channel energy into photosynthesis.

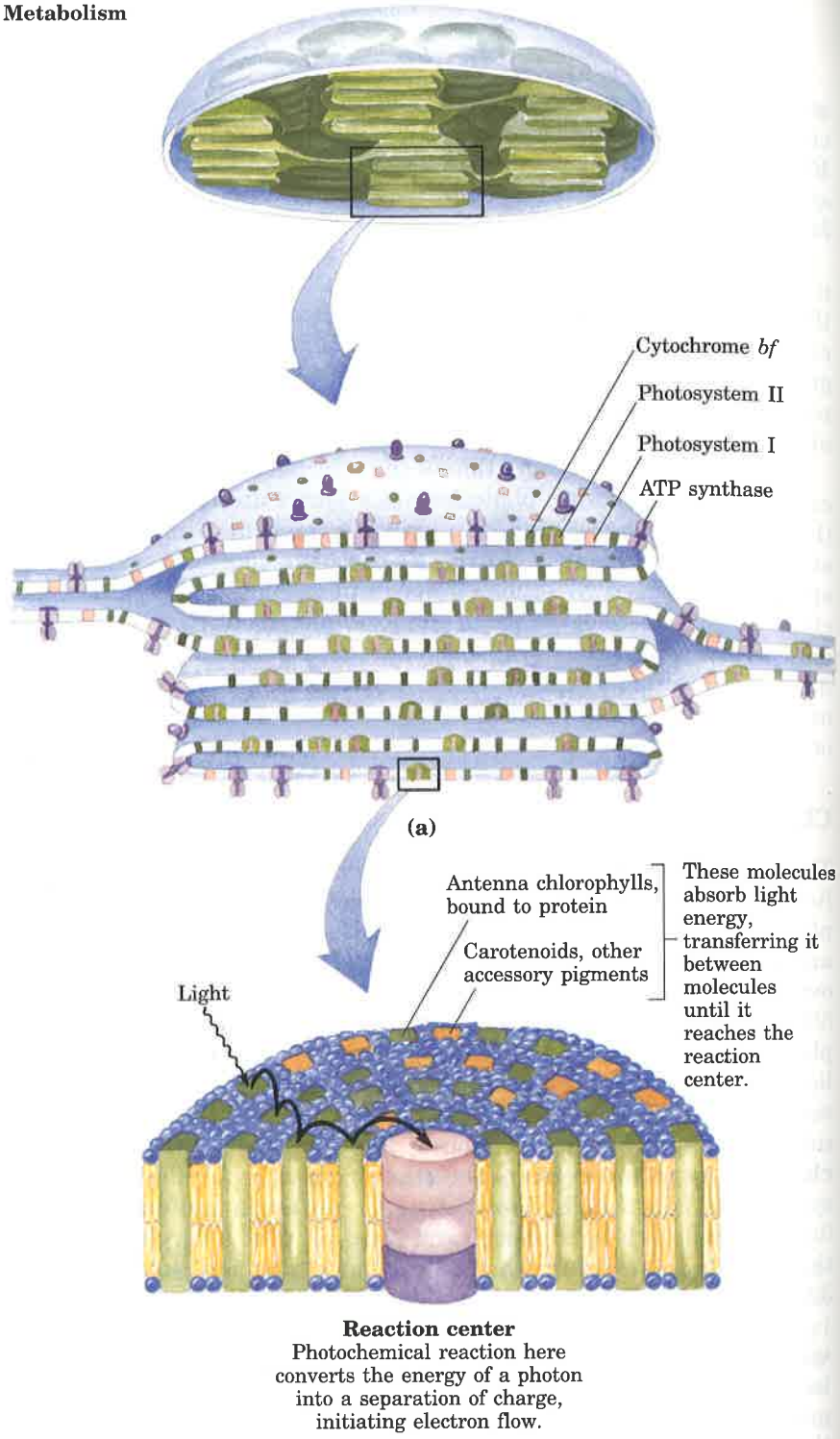



**Figure 18–39** Organization of the photosystem components in the thylakoid membrane. **(a)** The distribution of photosystems I and II, ATP synthase, and the cytochrome *bf* complex in the thylakoid membranes is not random. Photosystem I and ATP synthase are almost completely excluded from the regions with tightly stacked membranes, whereas photosystem II and the cytochrome *bf* complex are enriched in these regions of tight packing. This separation of photosystems I and II prevents energy absorbed by photosystem II from being transferred directly to photosystem I, and also places photosystem I in the regions most accessible to $NADP^+$ from the stroma. **(b)** An enlargement of a photosystem showing the reaction center, antenna chlorophylls, and accessory pigments. Cytochrome *bf* and ATP synthase of chloroplasts are described later in this chapter.

acceptor that is part of the electron transfer chain of the chloroplast, leaving the excited chlorophyll molecule with an empty orbital (an "electron hole") (step ④). The electron acceptor thus acquires a negative charge. The electron lost by the reaction-center chlorophyll is replaced by an electron from a neighboring electron donor molecule (step ⑤), which becomes positively charged. In this way, *excitation by light causes electric charge separation and initiates an oxidation–reduction chain.* Coupled to the light-dependent electron flow along this chain are processes that generate ATP and NADPH.

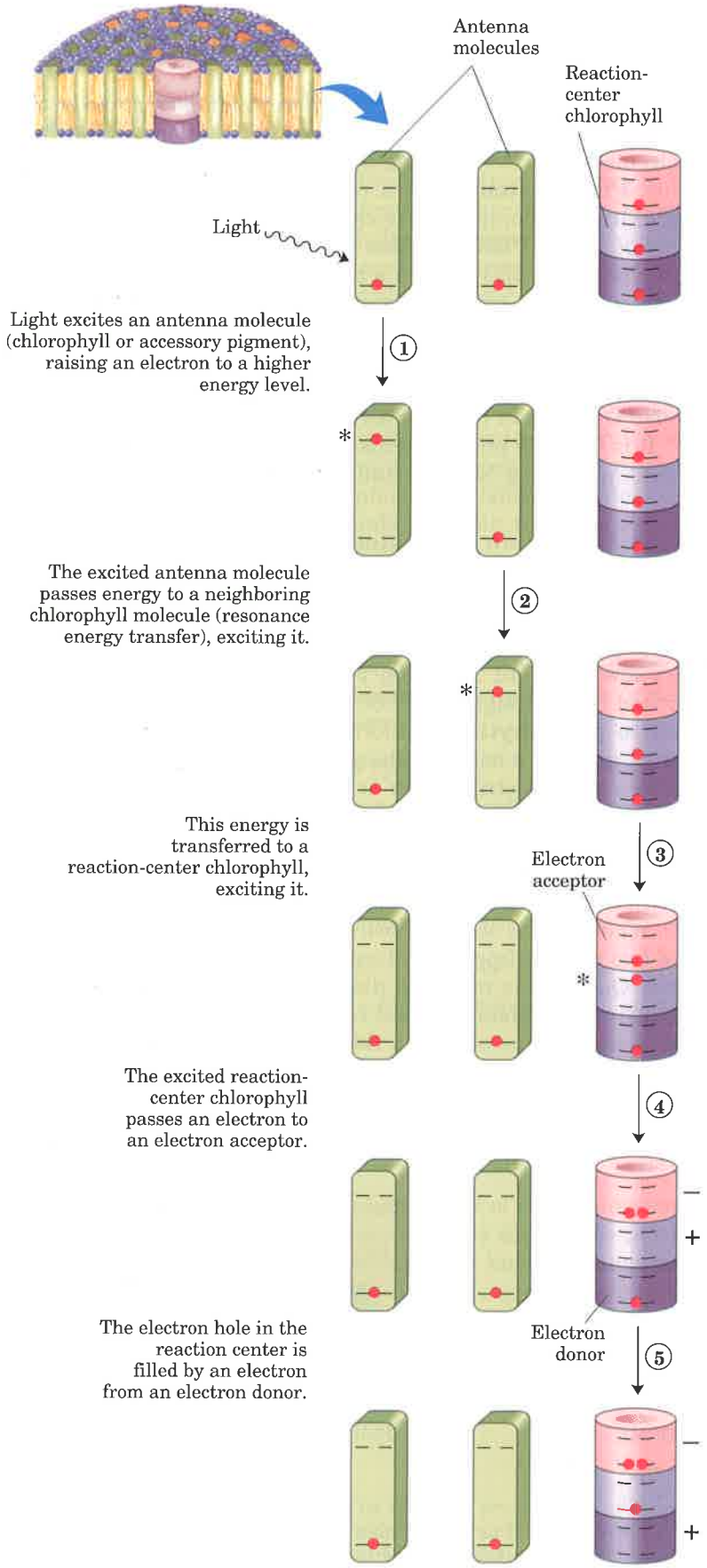



**Figure 18–40** A generalized scheme showing the conversion of energy from an absorbed photon into separation of charges at the photosystem reaction center. The steps are further described in the text. Note that step (2) may be repeated a number of times between successive antenna molecules until a reaction-center chlorophyll is reached. The asterisk (*) represents the excited state of an antenna molecule.

## Light-Driven Electron Flow

Thylakoid membranes have two different kinds of photosystems, each with its own type of photochemical reaction center and a set of antenna molecules. The two systems have distinct and complementary functions. **Photosystem I** has a reaction center designated **P700** and a high ratio of chlorophyll *a* to chlorophyll *b*. **Photosystem II,** with its reaction center **P680,** contains roughly equal amounts of chlorophyll *a* and *b* and may also contain a third type, chlorophyll *c*. The thylakoid membranes of a single spinach chloroplast have many hundreds of each kind of photosystem. All $O_2$-evolving photosynthetic cells—those of higher plants, algae, and cyanobacteria—contain both photosystems I and II; all other species of photosynthetic bacteria, which do not evolve $O_2$, contain only photosystem I.

It is between photosystems I and II that light-driven electron flow occurs, producing NADPH and a transmembrane proton gradient.

### Light Absorption by Photosystem II Initiates Charge Separation

How can light energy captured by chloroplasts induce electrons to flow energetically "uphill"? Excitation briefly creates a chemical species of very low standard reduction potential—an excellent electron donor. Excited P680, designated $P680^*$, within picoseconds transfers an electron to **pheophytin** (a chlorophyll-like accessory pigment lacking $Mg^{2+}$), giving it a negative charge (designated $Ph^-$) (Fig. 18–41a). With the loss of its electron, $P680^*$ is transformed into a radical cation, designated $P680^+$. Thus excitation, in creating $Ph^-$ and $P680^+$, causes charge separation. $Ph^-$ very rapidly passes its extra electron to a protein-bound **plastoquinone,** $Q_A$, which in turn passes its electron to another, more loosely bound quinone, $Q_B$ (Fig. 18–42; see also Fig. 18–44). When $Q_B$ has acquired two electrons in two such transfers from $Q_A$ and two protons from the solvent water, it is in its fully reduced quinol form, $Q_BH_2$. This molecule dissociates from its protein and diffuses away from the photochemical reaction center, carrying in its chemical bonds some of the energy of the photons that originally excited P680. The overall reaction initiated by light in photosystem II is therefore

$$4\ P680 + 4H^+ + 2Q_B + \text{light (4 photons)} \longrightarrow 4\ P680^+ + 2Q_BH_2 \quad (18\text{–}4)$$

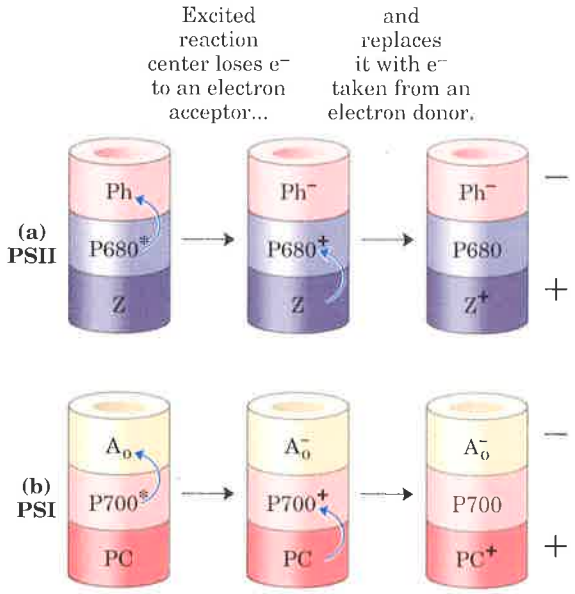



**Figure 18–41** Photochemical events following excitation of photosystems by light absorption. (The steps shown here are equivalent to steps ④ and ⑤ in Fig. 18–40.) **(a)** Photosystem II (PSII). Z represents a Tyr residue in the D1 protein of PSII; Ph, pheophytin. **(b)** Photosystem I (PSI). $A_0$ is a chlorophyll molecule near the reaction center of PSI; it accepts an electron from P700 to become the powerful reducing agent $A_0^-$. PC, plastocyanin.

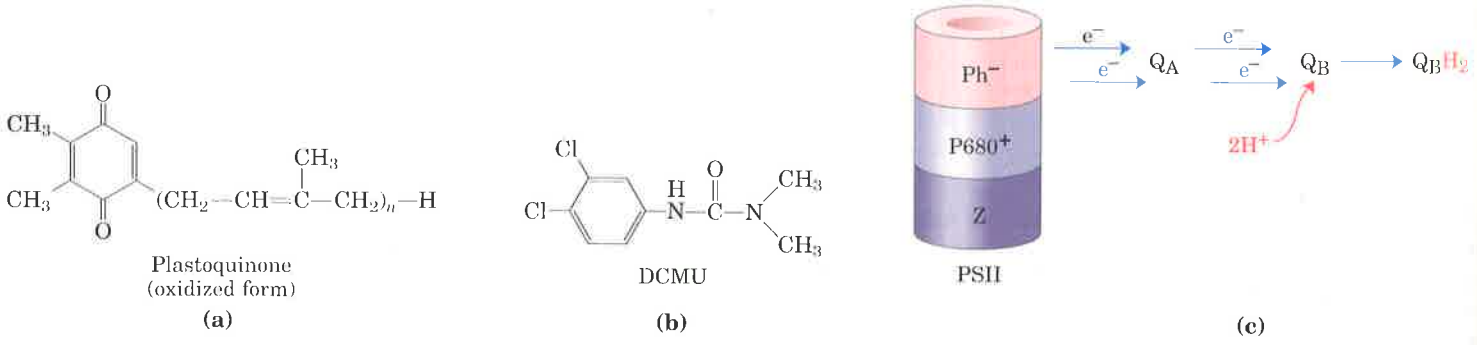



**Figure 18–42** **(a)** Plastoquinone ($Q_A$). **(b)** The herbicide DCMU (3-(3,4-dichlorophenyl)-1,1-dimethylurea), which displaces $Q_B$ from its binding site in photosystem II and blocks electron transfer from photosystem II to photosystem I. **(c)** The role of $Q_A$ and $Q_B$ in transferring electrons away from photosystem II. $Q_BH_2$ carries some of the energy of light absorbed by PSII.

Eventually, the electrons in $Q_BH_2$ are transferred through a chain of membrane-bound carriers to $NADP^+$, reducing it to NADPH and releasing $H^+$ (as described later). The potent herbicide DCMU (Fig. 18–42) competes with $Q_B$ for the $Q_B$ binding site in photosystem II, thus blocking photosynthetic electron transfer (Table 18–4).

In the meantime, $P680^+$ must acquire an electron to return to its ground state in preparation for the capture of another photon of energy (Fig. 18–41a). In principle, the required electron might come from any number of organic or inorganic compounds. Photosynthetic bacteria can use a variety of electron donors for this purpose—acetate, succinate, malate, or sulfide—depending on what is available in a particular ecological niche. About 3 billion years ago, evolution of primitive photosynthetic bacteria (the progenitors of the modern cyanobacteria) produced a photosystem capable of taking electrons from a donor that is always available—water. In this process two water molecules are split, yielding four electrons, four protons, and molecular oxygen: $2H_2O \longrightarrow 4H^+ + 4e^- + O_2$. A single photon of visible light does not possess enough energy to break the bonds in water; four photons are required in this photolytic cleavage reaction.

The four electrons abstracted from water do not pass directly to $P680^+$, which can only accept one electron at a time. Instead, a remarkable molecular device, the **water-splitting complex,** passes four electrons one at a time to $P680^+$. The immediate electron donor to $P680^+$ is a Tyr residue (often represented by the symbol Z) in protein D1 of the photosystem II reaction center:

$$4\ P680^+ + 4Z \longrightarrow 4\ P680 + 4Z^+ \qquad (18–5)$$

This Tyr residue regains its missing electron by oxidizing a cluster of four manganese ions in the water-splitting complex. With each single-electron transfer, this Mn cluster becomes more oxidized; four single-electron transfers, each corresponding to the absorption of one photon, produce a charge of +4 on the Mn complex (Fig. 18–43):

$$4Z^+ + [\text{Mn complex}]^0 \longrightarrow 4Z + [\text{Mn complex}]^{4+} \qquad (18–6)$$

In this state, the Mn complex can take four electrons from a pair of water molecules, releasing $4H^+$ and $O_2$:

$$[\text{Mn complex}]^{4+} + 2H_2O \longrightarrow [\text{Mn complex}]^0 + 4H^+ + O_2 \qquad (18–7)$$

The sum of Equations 18–4 through 18–7 is

$$2H_2O + 2Q_B + 4\ \text{photons} \longrightarrow O_2 + 2Q_BH_2 \qquad (18–8)$$

The water-splitting activity is an integral part of the photosystem II reaction center, and it has proved exceptionally difficult to purify. The detailed structure of the Mn cluster is not yet known. Manganese can exist in stable oxidation states from +2 to +7, so a cluster of four Mn ions can certainly donate or accept four electrons; the chemical details of this process, however, remain to be clarified.

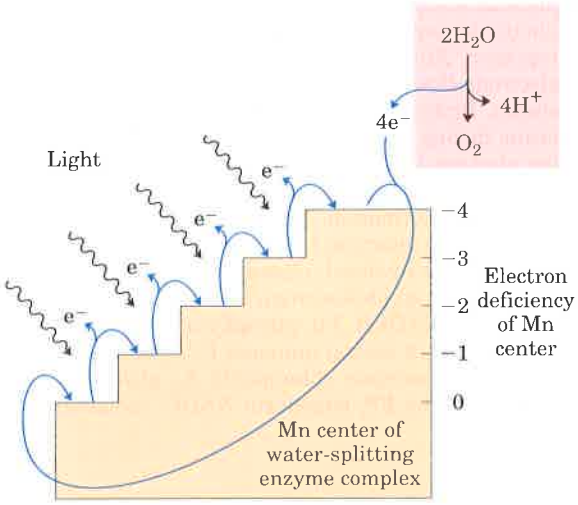



**Figure 18–43** The four-step process that produces a four-electron oxidizing agent, believed to be a complex of several Mn ions, in the water-splitting complex of photosystem II. The sequential absorption of four photons, each causing the loss of one electron from the Mn center, produces an oxidizing agent that can take four electrons from two molecules of water, producing $O_2$. The electrons lost from the Mn center pass one at a time to a Tyr residue ($Z^+$) in a reaction-center protein.

## Light Absorption by Photosystem I Creates a Powerful Reducing Agent

The photochemical events that follow excitation of photosystem I (P700) are formally similar to those in photosystem II (following the general scheme of Figure 18–40). Light is first captured by any one of about 200 chlorophyll *a* and *b* molecules, or additional accessory pigments, that serve as antennae, and the absorbed energy moves to P700 by resonance energy transfer. The excited reaction center $P700^*$ loses an electron to an acceptor, $A_0$ (believed to be a special form of chlorophyll, functionally analogous to the pheophytin of photosystem II), creating $A_0^-$ and $P700^+$ (Fig. 18–41b); again, excitation has resulted in charge separation at the photochemical reaction center. $P700^+$ is a strong oxidizing agent, which quickly acquires an electron from **plastocyanin,** a soluble Cu-containing electron transfer protein.

$A_0^-$ is an exceptionally strong reducing agent, which passes its electron through a chain of carriers that leads to $NADP^+$ (Fig. 18–44). First, **phylloquinone** ($A_1$) accepts an electron from $A_0^-$ and passes it on to an iron–sulfur protein. From here, the electron moves to **ferredoxin** (Fd), another iron–sulfur protein loosely associated with the thylakoid membrane. Spinach ferredoxin ($M_r$ 10,700), which has been isolated and crystallized, contains a 2Fe–2S center (Fig. 18–5). The Fe atoms of ferredoxin transfer electrons via one-electron $Fe^{2+}$ to $Fe^{3+}$ valence changes.

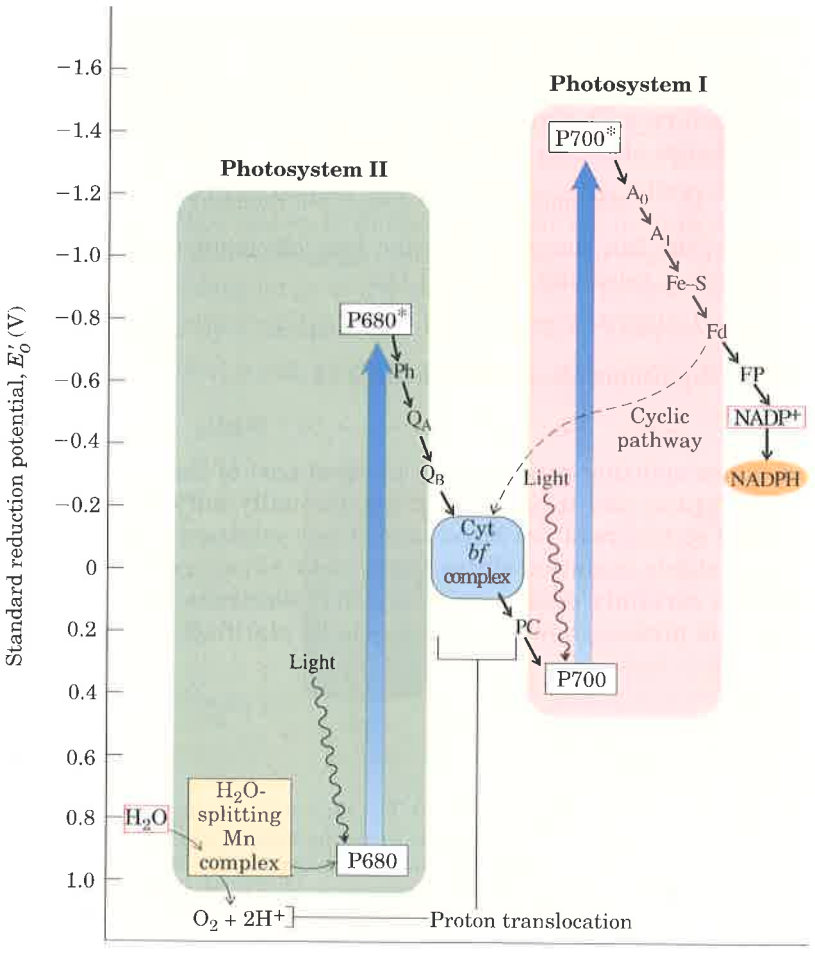



**Figure 18–44** The integration of photosystems I and II. This "Z scheme" shows the pathway of electron transfer from $H_2O$ (lower left) to $NADP^+$ (upper right) in noncyclic photosynthesis. The position on the vertical scale of each electron carrier reflects its standard reduction potential. To raise the energy of electrons derived from $H_2O$ to the energy level required to reduce $NADP^+$ to NADPH, each electron must be "lifted" twice (heavy arrows) by photons absorbed in photosystems I and II. One photon is required per electron boosted in each photosystem. After each excitation, the high-energy electrons flow "downhill" via the carrier chains shown. Protons move across the thylakoid membrane during the water-splitting reaction and during electron transfer through the cytochrome *bf* complex, producing the proton gradient that is central to ATP formation. The dashed arrow is the path of cyclic electron transfer, in which only photosystem I is involved; electrons return via the cyclic pathway to photosystem I, instead of reducing $NADP^+$ to NADPH. Ph, pheophytin; $Q_A$, plastoquinone; $Q_B$, a second quinone; PC, plastocyanin; $A_o$, electron acceptor chlorophyll; $A_1$, phylloquinone; Fd, ferredoxin; FP, ferredoxin-$NADP^+$ oxidoreductase.

The fourth electron carrier in the chain is a flavoprotein called **ferredoxin-NADP$^+$ oxidoreductase.** It transfers electrons from reduced ferredoxin ($Fd^{2+}_{red}$) to $NADP^+$, reducing the latter to NADPH:

$$2Fd^{2+}_{red} + 2H^+ + NADP^+ \longrightarrow 2Fd^{3+}_{ox} + NADPH + H^+$$

## Photosystems I and II Cooperate to Carry Electrons from $H_2O$ to $NADP^+$

Early studies by Robert Emerson established that maximum rates of photosynthesis, measured as $O_2$ evolution, required light of at least two wavelengths, now known to excite photosystems I and II. The two photosystems must function together in the $O_2$-evolving light reactions of photosynthesis. The diagram in Figure 18–44, often called the **Z scheme** because of its overall form, outlines the pathway of electron flow between the two photosystems as well as the energy relationships in the light reactions.

When photons are absorbed by photosystem I, electrons are expelled from the reaction center and flow down a chain of electron carriers to $NADP^+$ to reduce it to NADPH. $P700^+$, transiently electron-deficient, accepts an electron expelled by illumination of photosystem II, which arrives via a second connecting chain of electron carriers. This leaves an "electron hole" in photosystem II, which is filled by electrons from $H_2O$. We have described how water is split to yield: (1) electrons, which are donated to the electron-deficient photosystem II; (2) $H^+$ ions (protons), which are released inside the thylakoid lumen; and (3) $O_2$, which is released into the gas phase. The Z scheme thus describes the complete route by which electrons flow from $H_2O$ to $NADP^+$ according to the equation

$$2H_2O + 2NADP^+ + 8 \text{ photons} \longrightarrow O_2 + 2NADPH + 2H^+$$

For each electron transferred from $H_2O$ to $NADP^+$, two photons are absorbed, one by each photosystem. To form one molecule of $O_2$, which requires transfer of four electrons from two $H_2O$ to two $NADP^+$, a total of eight photons must be absorbed, four by each photosystem.

## The Cytochrome *bf* Complex Links Photosystems II and I

Electrons stored in $Q_BH_2$ as a result of the excitation of P680 in photosystem II (Fig. 18–42c) are carried to P700 of photosystem I via an assembly of several integral membrane proteins known as the **cytochrome *bf* complex** and the soluble protein plastocyanin (Fig. 18–44). The purified cytochrome *bf* complex contains a *b*-type cytochrome with two heme groups (cytochrome $b_{563}$), an iron–sulfur protein ($M_r$ 20,000), and cytochrome *f* (named for the Latin *frons*, meaning "leaf"), also called cytochrome $c_{552}$. Electrons flow through the cytochrome *bf* complex from $Q_BH_2$ to cytochrome *f*; the detailed path is uncertain. Cytochrome *f* passes its electron to plastocyanin, the donor for P700 reduction (Figs. 18–41b, 18–44).

The cytochrome *bf* complex of plants is remarkably similar to the cytochrome $bc_1$ complex (Complex III) of the mitochondrial electron transfer chain, and it carries out a similar function. Both complexes convey electrons from a reduced quinone, a mobile, lipid-soluble carrier of two electrons (UQ in mitochondria, $Q_B$ in chloroplasts), to a water-soluble protein that carries one electron (cytochrome *c* in mitochondria, plastocyanin in chloroplasts). As in mitochondria, the function of

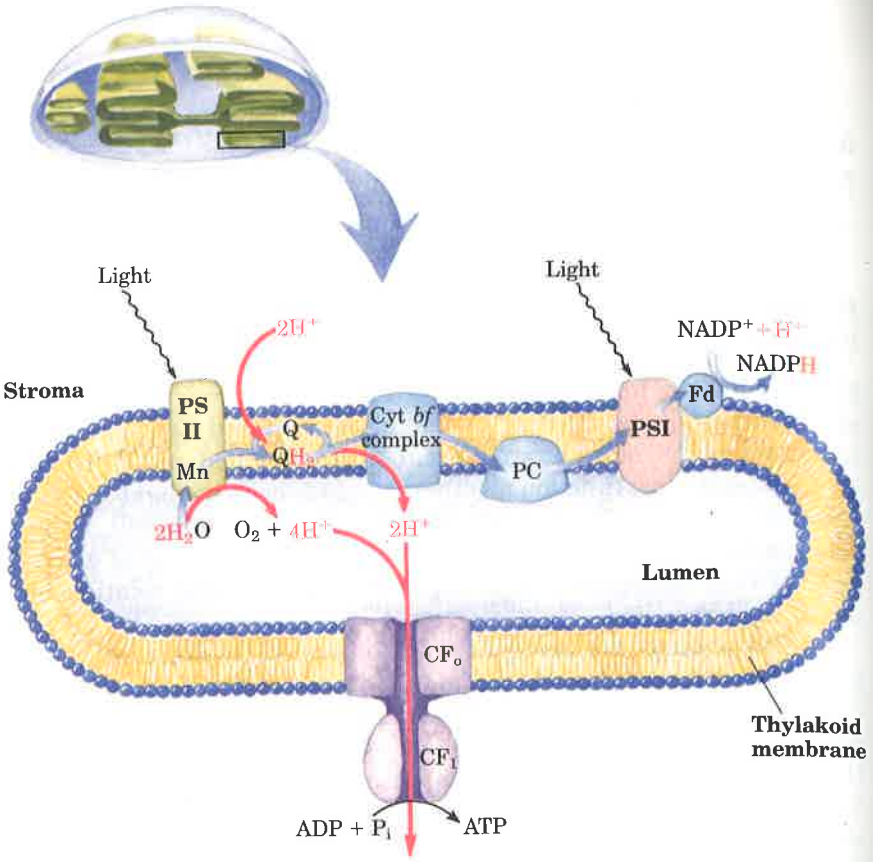



**Figure 18–45** Proton and electron circuits in chloroplast thylakoids. Electrons (blue arrows) move from $H_2O$ through photosystem II, the intermediate chain of carriers, photosystem I, and finally to $NADP^+$. Protons (red arrows) are pumped into the thylakoid lumen by the flow of electrons through the chain of carriers between photosystem II and photosystem I, and reenter the stroma through proton channels formed by the $F_o$ portion of the ATP synthase, designated $CF_o$ in the chloroplast enzyme. The $F_1$ subunit ($CF_1$) catalyzes synthesis of ATP.

this complex involves a Q cycle (see Fig. 18–10) in which electrons pass from $Q_BH_2$ to cytochrome *b* one at a time. As in the mitochondrial Complex III, this cycle results in the pumping of protons across the membrane; in chloroplasts, the direction of proton movement is from the stromal compartment to the thylakoid lumen. The result is the production of a proton gradient across the thylakoid membrane as electrons pass from photosystem II to photosystem I (Fig. 18–45). Because the volume of the flattened thylakoid lumen is small, the influx of a small number of protons has a relatively large effect on lumenal pH. The measured difference in pH between the stroma (pH 8) and the thylakoid lumen (pH 4.5) represents a 3,000-fold difference in proton concentration—a powerful driving force for ATP synthesis.

Daniel Arnon

## Coupling ATP Synthesis to Light-Driven Electron Flow

We have now seen how one of the two energy-rich products formed in the light reactions, NADPH, is generated by photosynthetic electron transfer from $H_2O$ to $NADP^+$. What about the other energy-rich product, ATP?

In 1954 Daniel Arnon and his colleagues discovered that ATP is generated from ADP and $P_i$ during photosynthetic electron transfer in illuminated spinach chloroplasts. Support for these findings came from the work of Albert Frenkel who detected light-dependent ATP production in membranous pigment-containing structures called **chromatophores,** derived from photosynthetic bacteria. They concluded that some of the light energy captured by the photosynthetic systems of these organisms is transformed into the phosphate bond energy of ATP. This process is called **photophosphorylation** or **photosyn-**

**thetic phosphorylation,** to distinguish it from oxidative phosphorylation in respiring mitochondria.

Recall that oxidative phosphorylation of ADP to ATP in mitochondria occurs at the expense of the free energy released as high-energy electrons flow downhill along the electron transfer chain from substrates to $O_2$. In a similar way, photophosphorylation of ADP to ATP is coupled to the energy released as high-energy electrons flow down the photosynthetic electron transfer chain from excited photosystem II to the electron-deficient photosystem I. The direct effect of electron flow is the formation of a proton gradient, which then provides the energy for ATP synthesis by an ATP synthase.

ATP synthesis in chloroplasts can be coupled to two types of electron flow—cyclic and noncyclic—as we shall see. We will also turn our attention to photophosphorylation in organisms other than green plants, and to the possible bacterial origins of chloroplasts. The chapter concludes with the development of an overall equation for photosynthesis in plants.

## A Proton Gradient Couples Electron Flow and Phosphorylation

Several properties of photosynthetic electron transfer and photophosphorylation in chloroplasts show a role for a proton gradient as in mitochondrial oxidative phosphorylation: (1) the reaction centers, electron carriers, and ATP-forming enzymes are located in a membrane—the thylakoid membrane; (2) photophosphorylation requires intact thylakoid membranes; (3) the thylakoid membrane is impermeable to protons; (4) photophosphorylation can be uncoupled from electron flow by reagents that promote the passage of protons through the thylakoid membrane; (5) photophosphorylation can be blocked by venturicidin and similar agents that inhibit the formation of ATP from ADP and $P_i$ by ATP synthase of mitochondria (see Fig. 18–13); and (6) ATP synthesis is catalyzed by $F_oF_1$ complexes, located on the outer surface of the thylakoid membranes, that are very similar in structure and function to the $F_oF_1$ complexes of mitochondria.

Electron-transferring molecules in the connecting chain between photosystem II and photosystem I are oriented asymmetrically in the thylakoid membrane, so that photoinduced electron flow results in the net movement of protons across the membrane, from the *outside* of the thylakoid membrane to the inner compartment (Fig. 18–45).

In 1966 André Jagendorf showed that a pH gradient across the thylakoid membrane (alkaline outside) could furnish the driving force to generate ATP. Jagendorf's early observations provided some of the most important experimental evidence in support of Mitchell's chemiosmotic hypothesis. In the dark, he soaked chloroplasts in a pH 4 buffer, which slowly penetrated into the inner compartment of the thylakoids, lowering their internal pH. He added ADP and $P_i$ to the dark suspension of chloroplasts and then suddenly raised the pH of the outer medium to 8, momentarily creating a large pH gradient across the membrane. As protons moved out of the thylakoids into the medium, ATP was generated from ADP and $P_i$. Because the formation of ATP occurred in the dark (with no input of energy from light), this experiment showed that a pH gradient across the membrane is a high-energy state that can, as in mitochondrial oxidative phosphorylation, mediate the transduction of energy from electron transfer into the chemical energy of ATP.

André Jagendorf

The stoichiometry for this process (protons transported per electron) is not well established. Electron transfer between photosystems II and I through the cytochrome *bf* complex contributes to the proton gradient an amount of proton-motive force roughly equivalent to one to two ATP formed per pair of electrons.

### The ATP Synthase of Chloroplasts Is Like That of Mitochondria

The enzyme responsible for ATP synthesis in chloroplasts is a large complex with two functional components, $CF_o$ and $CF_1$ (the C denoting chloroplast origin). $CF_o$ is a transmembrane proton pore composed of several integral membrane proteins and is homologous with mitochondrial $F_o$. $CF_1$ is a peripheral membrane protein complex very similar in subunit composition, structure, and function to mitochondrial $F_1$ (Table 18–8). Together these proteins constitute the ATP synthase of chloroplasts. (Bacteria also contain ATP synthases remarkably similar in structure and function to those of chloroplasts and mitochondria (Table 18–8).)

**Table 18–8 Equivalent subunits in ATP synthase of mitochondria, chloroplasts, and bacteria (*E. coli*)***

| Portion of ATP synthase | Mitochondria | | Chloroplasts | | *E. coli* | |
|---|---|---|---|---|---|---|
| | Subunit | Number | Subunit | Number | Subunit | Number |
| $F_1$ | $\alpha$ | 3 | $\alpha$ | 3 | $\alpha$ | 3 |
| | $\beta$ | 3 | $\beta$ | 3 | $\beta$ | 3 |
| | $\gamma$ | 1 | $\gamma$ | 1 | $\gamma$ | 1 |
| | OSCP | 1 | $\delta$ | 1 | $\delta$ | 1 |
| | $\delta$ | 1 | $\epsilon$ | 1 | $\epsilon$ | 1 |
| | $\epsilon$ | 1 | — | — | — | — |
| $F_o$ | a | 1 | a (IV) | 2 | a | 1 |
| | b | 1 | b and b′ (I and II) | 1 + 1 | b | 2 |
| | c | 6–12 | c (III) | 6–12 | c | 10–12 |

*Source:* Data primarily from Walker, J.E., Lutter, R., Dupuis, A., & Runswick, M.J. (1991) Identification of the subunits of $F_1F_o$-ATPase from bovine heart mitochondria. *Biochemistry* **30,** 5369–5378.

* Subunits on the same horizontal line are structurally related and are believed to be functionally homologous. Chloroplasts contain two nonidentical subunits, b and b′, that are together the homologs of mitochondrial b; *E. coli* has two identical b subunits. OSCP is oligomycin sensitivity-conferring protein. Alternative nomenclatures for the four chloroplast $F_o$ subunits are shown in parentheses. In addition to the subunits shown here, the mitochondrial enzyme complex contains at least five more proteins with no apparent homologs in the chloroplast or bacterium: F6, inhibitor protein, A6L, d, and e. These subunits are released as soluble proteins when $F_1$ and $F_o$ are dissociated, but they are believed to be part of the intact, functioning complex in mitochondria.

Electron microscopy of sectioned chloroplasts shows ATP synthase complexes as knoblike projections on the *outside* (stromal) surface of thylakoid membranes; these correspond to the ATP synthase complexes seen to project on the *inside* (matrix) surface of the inner mitochondrial membrane (see Fig. 18–15). Thus both the orientation of the ATP synthase and the direction of proton pumping in chloroplasts are opposite to those in mitochondria. In both cases, the $F_1$ portion of ATP synthase is located on the more alkaline side of the membrane through which protons flow down their concentration gradient; the direction of proton flow relative to $F_1$ is the same in both cases (Fig. 18–46).

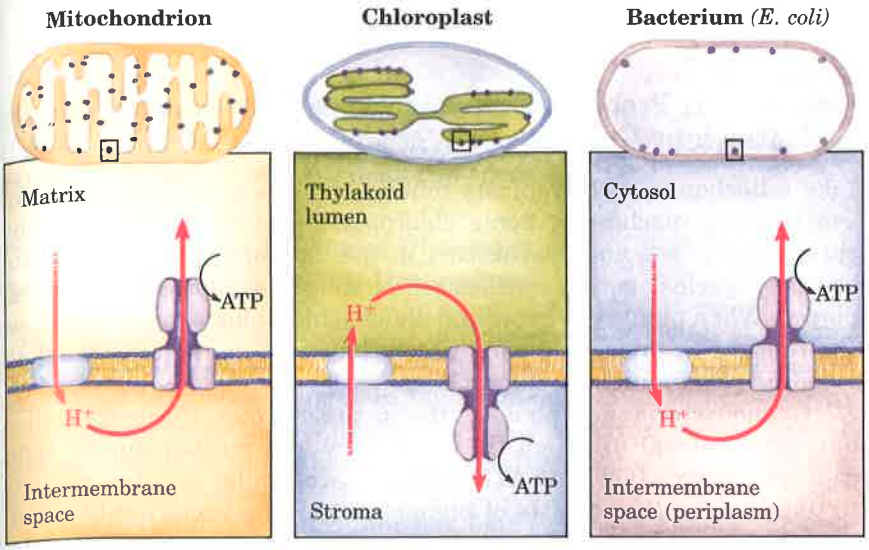



**Figure 18–46** Comparison of the topology of proton movement and ATP synthase orientation in the membranes of mitochondria, chloroplasts, and the bacterium *E. coli*. In each case, the orientation of the proton gradient relative to the ATP synthase is the same.

The mechanism of chloroplast ATP synthase is also believed to be essentially identical to that of its mitochondrial analog; ADP and $P_i$ readily condense to form ATP on the enzyme surface, but the release of this enzyme-bound ATP requires a proton-motive force (see Fig. 18–23). As for the mitochondrial ATP synthase, the details of this mechanism remain to be determined.

## Cyclic Electron Flow Produces ATP but Not NADPH or $O_2$

There is an alternative path of light-induced electron flow that allows chloroplasts to vary the ratio of NADPH and ATP formed during illuminations; this is called **cyclic electron flow** to differentiate it from the normally unidirectional or **noncyclic electron flow** that proceeds from $H_2O$ to $NADP^+$, as we have discussed thus far. Cyclic electron flow involves only photosystem I (Fig. 18–44). Electrons passed from P700 to ferredoxin do not continue to $NADP^+$, but move back through the cytochrome *bf* complex to plastocyanin. Plastocyanin donates electrons to P700, the illumination of which promotes electron transfer to ferredoxin. Thus illumination of photosystem I can cause electrons to cycle continuously out of the reaction center of photosystem I and back into it, each electron being propelled around the cycle by the energy yielded by absorption of one photon. Cyclic electron flow is not accompanied by net formation of NADPH or the evolution of $O_2$. However, it is accompanied by proton pumping and by the phosphorylation of ADP to ATP, referred to as **cyclic photophosphorylation.** The overall reaction equation for cyclic electron flow and photophosphorylation is simply

$$ADP + P_i \xrightarrow{\text{light}} ATP + H_2O$$

Cyclic electron flow and photophosphorylation are believed to occur when the plant cell is already amply supplied with reducing power in the form of NADPH but requires additional ATP for other metabolic needs. By regulating the partitioning of electrons between $NADP^+$ reduction and cyclic photophosphorylation, a plant adjusts the ratio of NADPH and ATP produced in the light reactions to match the needs for these products in the carbon fixation reactions and in other energy-requiring processes.

## Chloroplasts Probably Evolved from Endosymbiotic Cyanobacteria

Like mitochondria, chloroplasts contain their own DNA and protein-synthesizing machinery. Some chloroplast proteins are encoded by chloroplast genes and synthesized in the chloroplast; others are encoded by nuclear genes, synthesized outside the chloroplast, and imported. When plant cells grow and divide, chloroplasts give rise to new chloroplasts by division, during which their DNA is replicated and divided between daughter chloroplasts.

Cyanobacteria are photosynthetic prokaryotes (formerly called blue-green algae) in which the machinery and mechanism for light capture, electron flow, and ATP synthesis are similar in many respects to those in the chloroplasts of higher plants. The bacteriumlike division of chloroplasts, and their similarities to cyanobacteria in photosynthetic components and mechanisms, support the hypothesis that chloroplasts arose during evolution from endosymbiotic prokaryotes that also gave rise to modern cyanobacteria (see Fig. 2–17). Studies of photosynthesis in lower eukaryotes and prokaryotes may therefore yield insight into photosynthetic mechanisms that is generalizable to the higher plants.

## Diverse Photosynthetic Organisms Use Hydrogen Donors Other Than $H_2O$

Photosynthesis occurs not only in green plants but in lower eukaryotic organisms such as algae, euglenoids, dinoflagellates, and diatoms, and also in certain prokaryotes. The photosynthetic prokaryotes include the cyanobacteria, the green sulfur bacteria found in mountain lakes, the purple bacteria common in the ocean, and the purple sulfur bacteria of sulfur springs. The cyanobacteria, found in both fresh and salt waters, are perhaps the most versatile of photosynthetic organisms. Because they can also fix atmospheric nitrogen (Chapter 21), cyanobacteria are among the most self-sufficient organisms in the biosphere. At least half of the photosynthetic activity on earth occurs in the many different microorganisms that constitute the phytoplankton in oceans, rivers, and lakes.

With the exception of the cyanobacteria, which have an $O_2$-producing photosynthetic system resembling that of plants, photosynthetic bacteria do not produce $O_2$. Many are obligate anaerobes that cannot tolerate $O_2$. Some photosynthetic bacteria use inorganic compounds as hydrogen donors. For example, the green sulfur bacteria use hydrogen sulfide, according to the equation

$$2H_2S + CO_2 \xrightarrow{\text{light}} (CH_2O) + H_2O + 2S$$

These bacteria, instead of giving off molecular $O_2$, produce elemental sulfur as the oxidation product of $H_2S$. Other photosynthetic bacteria use organic compounds as hydrogen donors, for example, lactate:

$$2 \text{ Lactate} + CO_2 \xrightarrow{\text{light}} (CH_2O) + H_2O + 2 \text{ pyruvate}$$

Plant and bacterial photosynthesis are fundamentally similar processes despite the differences in the hydrogen donors they employ. This similarity becomes obvious when the equation of photosynthesis is written in a more general form:

$$2H_2D + CO_2 \xrightarrow{\text{light}} (CH_2O) + H_2O + 2D$$

in which $H_2D$ symbolizes a hydrogen donor and D is its oxidized form. $H_2D$ thus may be water, hydrogen sulfide, lactate, or some other organic compound, depending upon the species.

## The Structure of a Bacterial Photosystem Reaction Center Has Been Determined

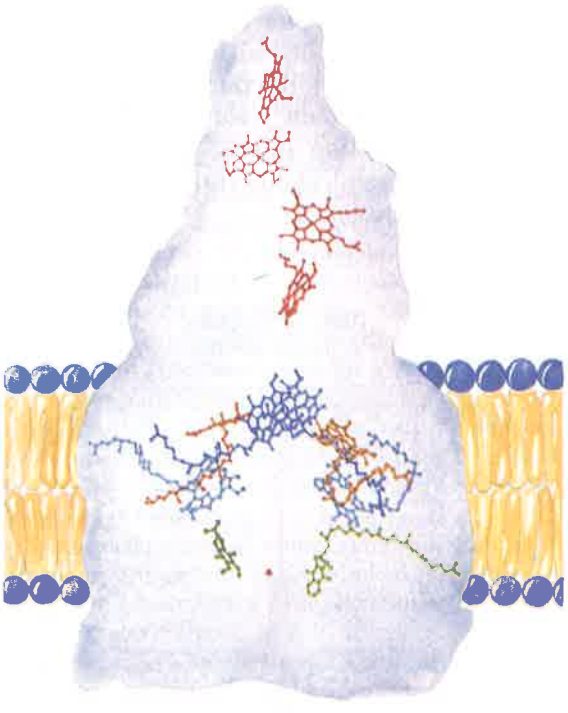
The three-dimensional structure of the photoreaction center of a photosynthetic bacterium, *Rhodopseudomonas viridis*, which is analogous in some ways to photosystem II in higher plants, is known from x-ray crystallographic studies (Fig. 18–47). This structure sheds light on how phototransduction takes place in the bacterium, and presumably in higher plants as well.

The bacterial reaction center has four types of proteins: a *c*-type cytochrome, two subunits (L and M) associated with bacteriochlorophyll, and a fourth protein, the H subunit. A single reaction center contains four hemes (cytochromes), four bacteriochlorophylls that are similar to the chlorophylls of chloroplasts, two bacteriopheophytins, one inorganic Fe, and a quinone. From a variety of physical studies, the extremely rapid sequence of events shown in Figure 18–47b has been deduced.

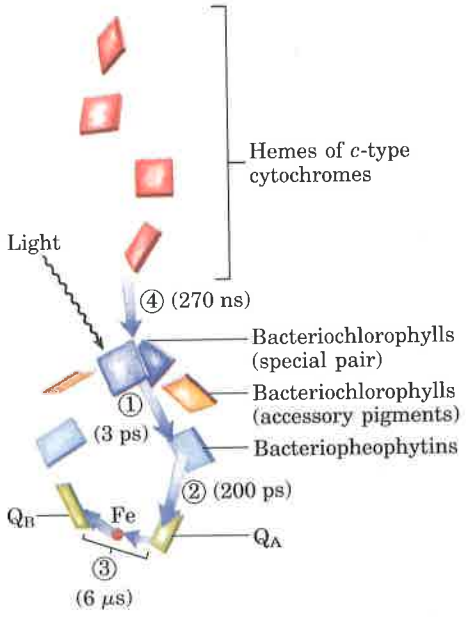
A pair of closely spaced bacteriochlorophylls (the "special pair") constitutes the site of the initial photochemistry in the bacterial reaction center. The "electron hole" that develops in the chlorophyll is filled with electrons from the *c*-type cytochrome, a role played by $H_2O$ in photosystem II of higher plants.

The bacteriochlorophylls, bacteriopheophytins, and quinone are held rigidly in a fixed orientation relative to each other by the proteins of the reaction center. The photochemical reactions among these components therefore take place in a virtually solid state, accounting for the high efficiency and rapidity of the reactions; nothing is left to chance collision or dependent upon random diffusion.

**Figure 18–47** The purple sulfur bacterium *Rhodopseudomonas viridis* performs light-driven ATP synthesis using machinery closely similar to photosystem II of higher plants, although the bacterium has no analog of photosystem I. **(a)** Superimposed on the structure of the reaction center proteins (see Fig. 10–11) are the prosthetic groups that participate in the photochemical events. There are four molecules of bacteriochlorophyll, two of which (the "special pair," dark blue) constitute the site of the first photochemical changes after light absorption. The other two bacteriochlorophylls (orange) are called accessory pigments; their role in the photochemical events is not well understood. Two bacterial pheophytins (light blue) lie beneath the bacteriochlorophylls, and beneath them, two bacterial quinones, $Q_A$ and $Q_B$ (green), separated by a non-heme iron atom (red). Shown at the top of the figure are four heme groups (red) associated with *c*-type cytochromes of the reaction center. **(b)** The sequence of events that occurs upon excitation of the special pair of bacteriochlorophylls and the time scale of the electron transfers (in parentheses). The excited special pair passes an electron to bacteriopheophytin ①, from which the electron moves rapidly to the tightly bound quinone $Q_A$ ②. This quinone much more slowly passes electrons to the diffusible quinone $Q_B$ through the non-heme iron atom ③. Meanwhile, the "electron hole" in the special pair is filled by an electron from one of the hemes of the four *c*-type cytochromes ④.



(a)




(b)

Although the reaction centers of plants have not yet been seen in such detail, the similarities between the bacterial and eukaryotic reaction centers suggest that the principles deduced from studies of the geometry and mechanism of the bacterial reaction center will also apply to the reaction centers of plants.

### Salt-Loving Bacteria Use Light Energy to Make ATP

The halophilic ("salt-loving") bacterium *Halobacterium halobium* conserves energy derived from absorbed sunlight by an interesting variation on the principle employed by true photosynthetic organisms. These unusual bacteria live only in brine ponds and salt lakes (Great Salt Lake and the Dead Sea, for example), where the high salt concentration results from water loss by evaporation; indeed, they cannot live in NaCl concentrations lower than 3 M. Halobacteria are aerobes and normally use $O_2$ to oxidize organic fuel molecules. However, the solubility of $O_2$ is so low in brine ponds, in which the NaCl concentration may exceed 4 M, that these bacteria must sometimes call on another source of energy, namely sunlight. The plasma membrane of *H. halobium* contains patches of light-absorbing pigments, called **purple patches.** These patches are made up of closely packed molecules of the protein **bacteriorhodopsin** (see Fig. 10–10), which contains retinal (vitamin A aldehyde; see Fig. 9–18) as a prosthetic group. When the cells are illuminated, the bacteriorhodopsin molecules are excited by an absorbed photon. As the excited molecules revert to their initial ground state, an induced conformational change results in the release of protons outside the cell, forming an acid-outside pH gradient across the plasma membrane. Protons tend to diffuse back into the cell through an ATP synthase complex in the membrane, very similar to that of mitochondria and of chloroplasts, supplying the energy for ATP synthesis (Fig. 18–48). Thus halobacteria can use light to supplement the ATP synthesized by oxidative phosphorylation with the $O_2$ that is available. However, halobacteria do not evolve $O_2$, nor do they carry out photoreduction of $NADP^+$; their phototransducing machinery is therefore much simpler than that of cyanobacteria or higher plants.

Bacteriorhodopsin, with only 247 amino acid residues, is the simplest light-driven proton pump known. The determination of its molecular structure should yield important insights into light-dependent energy transduction and the action of the proton pumps that function in respiration and photosynthesis.

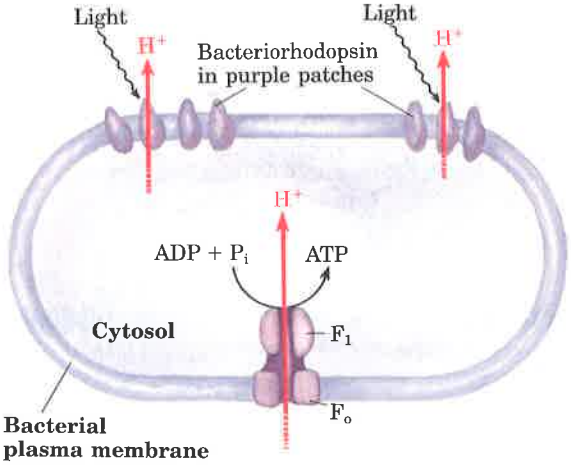



**Figure 18–48** Light-driven proton currents in *Halobacterium halobium* provide the proton-motive force for ATP synthesis. Illumination of the phototransducing protein bacteriorhodopsin results in outward proton movement, generating a proton-motive force. Reentry of protons through the $F_oF_1$ complex provides the energy for ATP synthesis.

### Photosynthesis Uses the Energy in Light Very Efficiently

The standard free-energy change for the synthesis of glucose from $CO_2$ and $H_2O$ during photosynthesis by the reaction

$$6CO_2 + 6H_2O \xrightarrow{\text{light}} C_6H_{12}O_6 + 6O_2 \qquad (18\text{–}9)$$

is 2,840 kJ/mol. (Recall that oxidation of glucose by the reverse of this equation proceeds with a *decrease* of 2,840 kJ/mol.) Now let us compare this energy requirement with the energy yielded by the light reactions of plant photosynthesis.

Recall that two photons must be absorbed, one by each photosystem, to cause flow of one electron from $H_2O$ to $NADP^+$. To generate one molecule of $O_2$, four electrons must be transferred. Therefore, production of six molecules of $O_2$ (as in Eqn 18–9) requires the absorption and use of 48 photons: (2 photons/$e^-$)(4$e^-$/$O_2$)(6$O_2$) = 48 photons. Because the energy of one einstein may range from 300 kJ at 400 nm to about 170 kJ at 700 nm (Fig. 18–35), anywhere from 8,160 to 14,400 kJ (depending upon the wavelength of the absorbed light) is required under standard conditions to make 1 mol of glucose "costing" 2,840 kJ.

In the next chapter we turn to a consideration of the carbon fixation reactions by which photosynthetic organisms use the ATP and NADPH produced in the light reactions to carry out the reduction of $CO_2$ to carbohydrates.

## Summary

Chemiosmotic theory provides the intellectual framework for understanding many biological energy transductions, including the processes of oxidative phosphorylation in mitochondria and photophosphorylation in chloroplasts. The mechanism of energy coupling is similar in both cases. The conservation of free energy involves the passage of electrons through a chain of membrane-bound oxidation–reduction (redox) carriers and the concomitant pumping of protons across the membrane, producing an electrochemical gradient, the protonmotive force. This force drives the synthesis of ATP by membrane-bound enzyme complexes through which protons flow back across the membrane, down their electrochemical gradient. Protonmotive force also drives other energy-requiring processes of cells.

In mitochondria, H atoms removed from substrates by the action of NAD-linked dehydrogenases donate their electrons to the respiratory (electron transfer) chain, which transfers them to molecular $O_2$, reducing it to $H_2O$. Shuttle systems convey reducing equivalents from cytosolic NADH to mitochondrial NADH. Reducing equivalents from all NAD-linked dehydrogenations are transferred to mitochondrial NADH dehydrogenase (Complex I), which contains FMN as its prosthetic group. They are then passed via a series of Fe–S centers to ubiquinone, which transfers the electrons to cytochrome *b*, the first carrier in Complex III. In this complex, electrons pass through two *b*-type cytochromes and cytochrome $c_1$ before reaching an Fe–S center. The Fe–S center passes electrons, one at a time, through cytochrome *c* and into Complex IV, cytochrome oxidase. This copper-containing enzyme, which also contains cytochromes *a* and $a_3$, accumulates electrons, then passes them to $O_2$, reducing it to $H_2O$.

There are alternative paths of entry of electrons into this chain of carriers. Succinate, for example, is oxidized by succinate dehydrogenase (Complex II), which contains a flavoprotein (with FAD) that passes electrons through several Fe–S centers and into the chain at the level of ubiquinone. Electrons derived from the oxidation of fatty acids pass into ubiquinone via the electron-transferring flavoprotein (ETFP).

The flow of electrons through Complexes I, III, and IV results in the pumping of protons across the mitochondrial inner membrane, making the matrix alkaline relative to the extramitochondrial space. This proton gradient provides the energy (proton-motive force) for ATP synthesis from ADP and $P_i$ by an inner-membrane protein complex, ATP synthase, also called $F_oF_1$ ATPase. The details of this ATP-synthesizing mechanism are still under investigation. Bacteria carry out oxidative phosphorylation by essentially the same mechanism, using electron carriers and an ATP synthase in the plasma membrane. Oxidative phosphorylation produces most of the ATP required by aerobic cells; it is regulated by cellular energy demands. In brown fat tissue, which is specialized for the production of metabolic heat, electron transfer is uncoupled from ATP synthesis; the energy of fatty acid oxidation is therefore dissipated as heat.

Photophosphorylation in the chloroplasts of green plants and in cyanobacteria also involves electron flow through a series of membrane-bound carriers. In the light reactions of plants, the absorption of a photon excites chlorophyll molecules and other (accessory) pigments that funnel the energy into reaction centers in the thylakoid membranes of chloroplasts. At the reaction centers, photoexcitation results in a charge separation that produces one chemical species that is a good electron donor (reducing agent) and another that is a good electron acceptor. In chloroplasts there are two different photoreaction centers, which function together. Photosystem I passes electrons from its excited reaction center, P700, through a series of carriers to ferredoxin, which then reduces $NADP^+$ to NADPH. The reaction center, P680, of photosystem II passes electrons to plastoquinone, reducing it to the quinol form. The electrons lost from P680 are replaced by electrons abstracted from $H_2O$ (hydrogen donors other than $H_2O$ are used in other organisms). This light-driven splitting of $H_2O$ is catalyzed by a Mn-containing protein complex; $O_2$ is produced. Reduced plastoquinone carries electrons from photosystem II to the cytochrome *bf* complex; these electrons pass to the soluble protein plastocyanin, and then to P700 to replace those lost during its photoexcitation. Electron flow through the cytochrome *bf* complex is accompanied by proton pumping across the thylakoid membrane, and the proton-motive force thus created drives ATP synthesis by a $CF_oCF_1$ complex closely similar to the $F_oF_1$ complex of mitochondria. This flow of electrons through photosystems II and I thus produces both NADPH and ATP. A second type of electron flow (cyclic flow) produces ATP only.

Both mitochondria and chloroplasts contain their own genomes and are believed to have originated from prokaryotic endosymbionts of early eukaryotic cells. Oxidative phosphorylation in aerobic bacteria and photophosphorylation in photosynthetic bacteria are closely similar, in machinery and mechanism, to the homologous processes in mitochondria and chloroplasts.

## Further Reading


***History and Background***

**Arnon, D.I.** (1984) The discovery of photosynthetic phosphorylation. *Trends Biochem. Sci.* **9,** 258–262.

**Harold, F.M.** (1986) *The Vital Force: A Study in Bioenergetics,* W.H. Freeman and Company, New York.

*A very readable synthesis of the principles of bioenergetics and their application to energy transductions.*

**Kalckar, H.M.** (1991) 50 years of biological research—from oxidative phosphorylation to energy requiring transport regulation. *Annu. Rev. Biochem.* **60,** 1–37.

*A delightful autobiographical account by one of the pioneers in the field.*

**Keilin, D.** (1966) *The History of Cell Respiration and Cytochrome,* Cambridge University Press, London.

*An authoritative and absorbing account of the discovery of cytochromes and of their roles in respiration, written by the discoverer of cytochromes.*

**Lehninger, A.L.** (1964) *The Mitochondrion: Molecular Basis of Structure and Function,* The Benjamin Co., Inc., New York.

*A classic description of early work on mitochondria.*

**Mitchell, P.** (1979) Keilin's respiratory chain concept and its chemiosmotic consequences. *Science* **206,** 1148–1159.

*The author's Nobel lecture, outlining the evolution of the chemiosmotic hypothesis.*

**Skulachev, V.P.** (1992) The laws of cell energetics. *Eur. J. Biochem.* **208,** 203–209.

*On the interconvertibility of ATP and ion gradients.*

**Slater, E.C.** (1987) The mechanism of the conservation of energy of biological oxidations. *Eur. J. Biochem.* **166,** 489–504.

*A clear and critical account of the evolution of the chemiosmotic model.*

**Staehelin, L.A. & Arntzen, C.J. (eds)** (1986) *Photosynthesis III: Photosynthetic Membranes and Light Harvesting Systems,* Encyclopedia of Plant Physiology, Vol. 19, Springer-Verlag, Berlin.

*Authoritative reviews of many aspects of photosynthesis.*

### *Respiratory Electron Flow*

**Babcock, G.T. & Wickström, M.** (1992) Oxygen activation and the conservation of energy in cell respiration. *Nature* **356,** 301–309.

*An advanced discussion of the reduction of water and pumping of protons by cytochrome oxidase.*

**Douce, R. & Neuburger, M.** (1989) The uniqueness of plant mitochondria. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40,** 371–414.

*A focus on the features of plant mitochondria that distinguish them from mitochondria of animal cells.*

**Hinkle, P.C. & McCarty, R.E.** (1978) How cells make ATP. *Sci. Am.* **238** (March), 104–123.

*Although not recent, this is an excellent, readable, and well-illustrated description of oxidative phosphorylation.*

**Lehninger, A.L., Reynafarje, B., Alexandre, A., & Villalobo, A.** (1980) Respiration-coupled $H^+$ ejection by mitochondria. *Ann. N. Y. Acad. Sci.* **341,** 585–592.

*The methods and problems in measuring proton efflux stoichiometry.*

**Malmström, B.G.** (1989) The mechanism of proton translocation in respiration and photosynthesis. *FEBS Lett.* **250,** 9–21.

*Comparative review of the electron-transferring complexes of mitochondria and chloroplasts.*

**Trumpower, B.L.** (1990) The protonmotive Q cycle: energy transduction by coupling of proton translocation to electron transfer by the cytochrome $bc_1$ complex. *J. Biol. Chem.* **265,** 11409–11412.

*Short, clear description of the Q cycle and electron flow through Complex III.*

### *Coupling ATP Synthesis to Respiratory Electron Flow*

**Boyer, P.D.** (1989) A perspective of the binding change mechanism for ATP synthesis. *FASEB J.* **3,** 2164–2178.

*An article on the historical development and current state of the binding-change model, by its principal architect.*

**Futai, M., Noumi, T., & Maeda, M.** (1987) Molecular biological studies on structure and mechanism of proton translocating ATPase ($H^+$-ATPase, $F_oF_1$). *Adv. Biophys.* **23,** 1–37.

*Insight into the mechanism of ATP synthase from studies of the genes that encode its subunits.*

**Pedersen, P.L. & Carafoli, E.** (1987) Ion motive ATPases. I. Ubiquity, properties, and significance to cell function. *Trends Biochem. Sci.* **12,** 145–150. II. Energy coupling and work output. *Trends Biochem. Sci.* **12,** 186–189.

*Two short reviews that place ATP synthase within the family of ATP-dependent proton pumps; include their general mechanisms.*

**Penefsky, H.S. & Cross, R.L.** (1991) Structure and mechanism of $F_oF_1$-type ATP synthases and ATPases. *Adv. Enzymol. Relat. Areas Mol. Bio.* **64,** 173–214.

*An advanced discussion.*

**Ricquier, D., Casteilla, L., & Bouillaud, F.** (1991) Molecular studies of the uncoupling protein. *FASEB J.* **5,** 2237–2242.

*A discussion of the protein and its role in thermogenesis.*

**Senior, A.E.** (1988) ATP synthesis by oxidative phosphorylation. *Physiol. Rev.* **68,** 177–231.

*An advanced but very clear review, with an emphasis on the mechanism of ATP synthase.*

### *Regulation of Mitochondrial Oxidative Phosphorylation*

**Brand, M.D. & Murphy, M.P.** (1987) Control of electron flux through the respiratory chain in mitochondria and cells. *Biol. Rev. Cambridge Phil. Soc.* **62,** 141–193.

*An advanced description of respiratory control.*

**Harris, D.A. & Das, A.M.** (1991) Control of mitochondrial ATP synthesis in the heart. *Biochem. J.* **280,** 561–573.

*An advanced discussion of regulation of the ATP synthase by $Ca^{2+}$ and other factors.*

### *Photosynthesis: Harvesting Light Energy*

**Green, B.R., Pichersky, E., & Kloppstech, K.** (1991) Chlorophyll *a/b*-binding proteins: an extended family. *Trends Biochem. Sci.* **16,** 181–186.
*An intermediate-level description of the proteins that orient chlorophyll molecules in chloroplasts.*

**Huber, R.** (1990) A structural basis of light energy and electron transfer in biology. *Eur. J. Biochem.* **187,** 283–305.
*The author's Nobel lecture, describing the physics and chemistry of phototransductions. An exceptionally clear and well-illustrated discussion, based on crystallographic studies of reaction centers.*

**Zuber, H.** (1986) Structure of light-harvesting antenna complexes of photosynthetic bacteria, cyanobacteria and red algae. *Trends Biochem. Sci.* **11,** 414–419.

### *Light-Driven Electron Flow*

**Andréasson, L.-E. & Vänngård, T.** (1988) Electron transport in photosystems I and II. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **39,** 379–411.
*An advanced description of the path of electron flow in chloroplasts, studied with spectroscopic techniques.*

**Blankenship, R.E. & Prince, R.C.** (1985) Excited-state redox potentials and the Z scheme of photosynthesis. *Trends Biochem. Sci.* **10,** 382–383.
*A concise and lucid statement of the redox properties of excited states.*

**Deisenhofer, J. & Michel, H.** (1991) Structures of bacterial photosynthetic reaction centers. *Annu. Rev. Cell Biol.* **7,** 1–23.
*The structure of the reaction center of purple bacteria, and implications for the function of bacterial and plant reaction centers.*

**Glazer, A.N. & Melis, A.** (1987) Photochemical reaction centers: structure, organization, and function. *Annu. Rev. Plant Physiol.* **38,** 11–45.
*An advanced description of the structure and function of reaction centers of green plants, cyanobacteria, and purple and green bacteria.*

**Golbeck, J.H.** (1992) Structure and function of photosystem I. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **43,** 293–324.

**Govindjee & Coleman, W.J.** (1990) How plants make oxygen. *Sci. Am.* **262** (February), 50–58.
*An exceptionally clear account of the water-splitting activity of photosystem II.*

**Nitschke, W. & Rutherford, A.W.** (1991) Photosynthetic reaction centres: variations on a common structural theme? *Trends Biochem. Sci.* **16,** 241–245.
*A comparison of the structure and function of photosystems I and II and the reaction centers of several photosynthetic bacteria.*

### *Coupling ATP Synthesis to Light-Driven Electron Flow*

**Cramer, W.A., Widger, W.R., Herrmann, R.G., & Trebst, A.** (1985) Topography and function of thylakoid membrane proteins. *Trends Biochem. Sci.* **10,** 125–129.

**Jagendorf, A.T.** (1967) Acid-base transitions and phosphorylation by chloroplasts. *Fed. Proc.* **26,** 1361–1369.
*The classic experiment establishing the ability of a proton gradient to drive ATP synthesis in the dark.*

**Youvan, D.C. & Marrs, B.L.** (1987) Molecular mechanisms of photosynthesis. *Sci. Am.* **256** (June), 42–48.
*An excellent description of the chemical basis for light reactions.*


## Problems

**1.** *Oxidation–Reduction Reactions* The NADH dehydrogenase complex of the mitochondrial respiratory chain promotes the following series of oxidation–reduction reactions, in which $Fe^{3+}$ and $Fe^{2+}$ represent the iron in iron–sulfur centers, UQ is ubiquinone, $UQH_2$ is ubiquinol, and E is the enzyme:

(1) $NADH + H^+ + E\text{–}FMN \rightarrow NAD^+ + E\text{–}FMNH_2$
(2) $E\text{–}FMNH_2 + 2Fe^{3+} \rightarrow E\text{–}FMN + 2Fe^{2+} + 2H^+$
(3) $2Fe^{2+} + 2H^+ + UQ \rightarrow 2Fe^{3+} + UQH_2$
*Sum:* $NADH + H^+ + UQ \rightarrow NAD^+ + UQH_2$

For each of the three reactions catalyzed by the NADH dehydrogenase complex, identify (a) the

electron donor, (b) the electron acceptor, (c) the conjugate redox pair, (d) the reducing agent, and (e) the oxidizing agent.

**2.** *Standard Reduction Potentials* The standard reduction potential of any redox couple is defined for the half-cell reaction (or half-reaction):

$$\text{Oxidizing agent} + n \text{ electrons} \longrightarrow \text{reducing agent}$$

The standard reduction potentials of the $NAD^+$/NADH and pyruvate/lactate redox pairs are $-0.320$ and $-0.185$ V, respectively.

(a) Which redox pair has the greater tendency to lose electrons? Explain.

(b) Which is the stronger oxidizing agent? Explain.

(c) Beginning with 1 M concentrations of each reactant and product at pH 7, in which direction will the following reaction proceed?

$$\text{Pyruvate} + \text{NADH} + H^+ \rightleftharpoons \text{lactate} + NAD^+$$

(d) What is the standard free-energy change, $\Delta G^{\circ\prime}$, at 25 °C for this reaction?

(e) What is the equilibrium constant for this reaction at 25 °C?

**3.** *Energy Span of the Respiratory Chain* Electron transfer in the mitochondrial respiratory chain may be represented by the net reaction equation

$$\text{NADH} + H^+ + \tfrac{1}{2}O_2 \rightleftharpoons H_2O + NAD^+$$

(a) Calculate the value of the change in standard reduction potential, $\Delta E'_0$, for the net reaction of mitochondrial electron transfer.

(b) Calculate the standard free-energy change, $\Delta G^{\circ\prime}$, for this reaction.

(c) How many ATP molecules could *theoretically* be generated per molecule of NADH oxidized by this reaction, given a standard free energy of ATP synthesis of 30.5 kJ/mol?

(d) How many ATP molecules could be synthesized under typical cellular conditions (see Box 13–2)?

**4.** *Use of FAD Rather Than $NAD^+$ in the Oxidation of Succinate* All the dehydrogenation steps in glycolysis and the citric acid cycle use $NAD^+$ ($E'_0$ for $NAD^+$/NADH = $-0.32$ V) as the electron acceptor except succinate dehydrogenase, which uses covalently bound FAD ($E'_0$ for FAD/$FADH_2$ in this enzyme = 0.05 V). Why is FAD a more appropriate electron acceptor than $NAD^+$ in the dehydrogenation of succinate? Give a possible explanation based on a comparison of the $E'_0$ values of the fumarate/succinate pair ($E'_0 = 0.03$), the $NAD^+$/NADH pair, and the succinate dehydrogenase FAD/$FADH_2$ pair.

**5.** *Degree of Reduction of Electron Carriers in the Respiratory Chain* The degree of reduction of each electron carrier in the respiratory chain is determined by the conditions existing in the mitochondrion. For example, when the supply of NADH and $O_2$ is abundant, the steady-state degree of reduction of the carriers decreases as electrons pass from the substrate to $O_2$. When electron transfer is blocked, the carriers before the block become more reduced while those beyond the block become more oxidized (Fig. 18–7). For each of the conditions below, predict the state of oxidation of each carrier in the respiratory chain (ubiquinone and cytochromes $b$, $c_1$, $c$, and $a + a_3$).

(a) Abundant supply of NADH and $O_2$ but cyanide added

(b) Abundant supply of NADH but $O_2$ exhausted

(c) Abundant supply of $O_2$ but NADH exhausted

(d) Abundant supply of NADH and $O_2$

**6.** *The Effect of Rotenone and Antimycin A on Electron Transfer* Rotenone, a toxic natural product from plants, strongly inhibits NADH dehydrogenase of insect and fish mitochondria. Antimycin A, a toxic antibiotic, strongly inhibits the oxidation of ubiquinol.

(a) Explain why rotenone ingestion is lethal to some insect and fish species.

(b) Explain why antimycin A is a poison.

(c) Assuming that rotenone and antimycin A are equally effective in blocking their respective sites in the electron transfer chain, which would be a more potent poison? Explain.

**7.** *Uncouplers of Oxidative Phosphorylation* In normal mitochondria the rate of electron transfer is tightly coupled to the demand for ATP. Thus when the rate of utilization of ATP is relatively low, the rate of electron transfer is also low. Conversely, when ATP is demanded at a high rate, electron transfer is rapid. Under such conditions of tight coupling, the number of ATP molecules produced per atom of oxygen consumed when NADH is the electron donor—known as the P/O ratio—is close to 3.

(a) Predict the effect of a relatively low and a relatively high concentration of an uncoupling agent on the rate of electron transfer and the P/O ratio.

(b) The ingestion of uncouplers causes profuse sweating and an increase in body temperature. Explain this phenomenon in molecular terms. What happens to the P/O ratio in the presence of uncouplers?

(c) The uncoupler 2,4-dinitrophenol was once prescribed as a weight-reducing drug. How can this agent, in principle, serve as a reducing aid? Such uncoupling agents are no longer prescribed because some deaths occurred following their use. Why can the ingestion of uncouplers lead to death?

**8.** *Mode of Action of Dicyclohexylcarbodiimide (DCCD)* When DCCD is added to a suspension of tightly coupled, actively respiring mitochondria, the rate of electron transfer (measured by $O_2$ consumption) and the rate of ATP production dramatically decrease. If a solution of 2,4-dinitrophenol is

now added to the inhibited mitochondrial preparation, $O_2$ consumption returns to normal but ATP production remains inhibited.

(a) What process in electron transfer or oxidative phosphorylation is affected by DCCD?

(b) Why does DCCD affect the $O_2$ consumption of mitochondria? Explain the effect of 2,4-dinitrophenol on the inhibited mitochondrial preparation.

(c) Which of the following inhibitors does DCCD most resemble in its action: antimycin A, rotenone, or oligomycin?

**9.** *The Malate–α-Ketoglutarate Transport System of Mitochondria* The inner mitochondrial membrane transport system that promotes the transport of malate and α-ketoglutarate across the membrane (Fig. 18–25) is inhibited by *n*-butylmalonate. Suppose *n*-butylmalonate is added to an aerobic suspension of kidney cells using glucose exclusively as fuel. Predict the effect of this inhibitor on

(a) Glycolysis
(b) Oxygen consumption
(c) Lactate formation
(d) ATP synthesis

**10.** *The Pasteur Effect* When $O_2$ is added to an anaerobic suspension of cells using glucose at a high rate, the rate of glucose consumption declines dramatically as the added $O_2$ is consumed. In addition, the accumulation of lactate ceases. This effect, first observed by Louis Pasteur in the 1860s, is characteristic of most cells capable of both aerobic and anaerobic utilization of glucose.

(a) Why does the accumulation of lactate cease after $O_2$ is added?

(b) Why does the presence of $O_2$ decrease the rate of glucose consumption?

(c) How does the onset of $O_2$ consumption slow down the rate of glucose consumption? Explain in terms of specific enzymes.

**11.** *How Many Protons in a Mitochondrion?* Electron transfer functions to translocate protons from the mitochondrial matrix to the external medium to establish a pH gradient across the inner membrane, the outside more acidic than the inside. The tendency of protons to diffuse from the outside into the matrix, where $[H^+]$ is lower, is the driving force for ATP synthesis via the ATP synthase. During oxidative phosphorylation by a suspension of mitochondria in a medium of pH 7.4, the internal pH of the matrix has been measured as 7.7.

(a) Calculate $[H^+]$ in the external medium and in the matrix under these conditions.

(b) What is the outside:inside ratio of $[H^+]$? Comment on the energy inherent in this concentration. (Hint: See p. 383, Eqn 13–5.)

(c) Calculate the number of protons in a respiring liver mitochondrion, assuming its inner matrix compartment is a sphere of diameter 1.5 μm.

(d) From these data would you think the pH gradient alone is sufficiently great to generate ATP?

(e) If not, can you suggest how the necessary energy for synthesis of ATP arises?

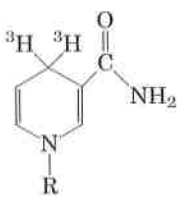
**12.** *Rate of ATP Turnover in Rat Heart Muscle* Rat heart muscle operating aerobically fills more than 90% of its ATP needs by oxidative phosphorylation. This tissue consumes $O_2$ at the rate of 10 μmol/min • g of tissue, with glucose as the fuel source.

(a) Calculate the rate at which this tissue consumes glucose and produces ATP.

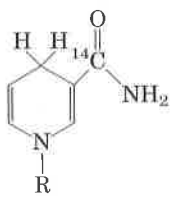
(b) If the steady-state concentration of ATP in rat heart muscle is 5 μmol/g of tissue, calculate the time required (in seconds) to completely turn over the cellular pool of ATP. What does this result indicate about the need for tight regulation of ATP production? (Note: Concentrations are expressed as micromoles per gram of muscle tissue because the tissue is mostly water.)

**13.** *Rate of ATP Breakdown in Flight Muscle* ATP production in the flight muscles of the fly *Lucilia sericata* results almost exclusively from oxidative phosphorylation. During flight, 187 ml of $O_2$/h • g of fly body weight is needed to maintain an ATP concentration of 7 μmol/g of flight muscle. Assuming that the flight muscles represent 20% of the weight of the fly, calculate the rate at which the flight-muscle ATP pool turns over. How long would the reservoir of ATP last in the absence of oxidative phosphorylation? Assume that reducing equivalents are transferred by the glycerol-3-phosphate shuttle and that $O_2$ is at 25 °C and 101.3 kPa (1 atm). (Note: Concentrations are expressed in micromoles per gram of flight muscle.)

**14.** *Transmembrane Movement of Reducing Equivalents* Under aerobic conditions, extramitochondrial NADH must be oxidized by the mitochondrial electron transfer chain. Consider a preparation of rat hepatocytes containing mitochondria and all the enzymes of the cytosol. If [4–$^3$H]NADH is introduced, radioactivity appears quickly in the mitochondrial matrix. However, if [7–$^{14}$C]NADH is introduced, no radioactivity appears in the matrix. What do these observations tell us about the oxidation of extramitochondrial NADH by the electron transfer chain?



[4-$^3$H]NADH



[7-$^{14}$C]NADH

**15.** *Photochemical Efficiency of Light at Different Wavelengths* The rate of photosynthesis, measured by $O_2$ production, is higher when a green plant is illuminated with light of wavelength 680 nm than with light of 700 nm. However, illumination by a combination of light of 680 nm and 700 nm gives a higher rate of photosynthesis than light of either wavelength alone. Explain.

**16.** *Role of $H_2S$ in Some Photosynthetic Bacteria* Illuminated purple sulfur bacteria carry out photosynthesis in the presence of $H_2O$ and $^{14}CO_2$, but only if $H_2S$ is added and $O_2$ is absent. During the course of photosynthesis, measured by formation of [$^{14}$C]carbohydrate, $H_2S$ is converted into elemental sulfur, but no $O_2$ is evolved. What is the role of the conversion of $H_2S$ into sulfur? Why is no $O_2$ evolved?

**17.** *Boosting the Reducing Power of Photosystem I by Light Absorption* When photosystem I absorbs red light at 700 nm, the standard reduction potential of P700 changes from 0.4 to about −1.2 V. What fraction of the absorbed light is trapped in the form of reducing power?

**18.** *Mode of Action of the Herbicide DCMU* When chloroplasts are treated with 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU, or Diuron), a potent herbicide, $O_2$ evolution and photophosphorylation cease. Oxygen evolution but not photophosphorylation can be restored by the addition of an external electron acceptor, or Hill reagent. How does this herbicide act as a weed killer? Suggest a location for the inhibitory site of this herbicide in the scheme shown in Figure 18–44. Explain.

**19.** *Bioenergetics of Photophosphorylation* The steady-state concentrations of ATP, ADP, and $P_i$ in isolated spinach chloroplasts under full illumination at pH 7.0 are 120, 6, and 700 $\mu$M, respectively.

(a) What is the free-energy requirement for the synthesis of 1 mol of ATP under these conditions?

(b) The energy for ATP synthesis is furnished by light-induced electron transfer in the chloroplasts. What is the minimum voltage drop necessary during the transfer of a pair of electrons to synthesize ATP under these conditions? (You may need to refer to p. 389, Eqn 13–8.)

**20.** *Equilibrium Constant for Water-Splitting Reactions* The coenzyme $NADP^+$ is the terminal electron acceptor in chloroplasts, according to the reaction

$$2H_2O + 2NADP^+ \longrightarrow 2NADPH + 2H^+ + O_2$$

Use the information in Table 18–2 to calculate the equilibrium constant at 25 °C for this reaction. (The relationship between $K'_{eq}$ and $\Delta G^{\circ\prime}$ is discussed on p. 368.) How can the chloroplast overcome this unfavorable equilibrium?

**21.** *Energetics of Phototransduction* During photosynthesis, eight photons of light must be absorbed (four by each photosystem) for every $O_2$ molecule produced:

$$2H_2O + 2NADP^+ + 8 \text{ photons} \longrightarrow 2NADPH + 2H^+ + O_2$$

Assuming that these photons have a wavelength of 700 nm (red) and that the absorption and utilization of light energy are 100% efficient, calculate the free-energy change for the process.

**22.** *Electron Transfer to a Hill Reagent* Isolated spinach chloroplasts evolve $O_2$ when illuminated in the presence of potassium ferricyanide (the Hill reagent), according to the equation

$$2H_2O + 4Fe^{3+} \longrightarrow O_2 + 4H^+ + 4Fe^{2+}$$

where $Fe^{3+}$ represents ferricyanide and $Fe^{2+}$, ferrocyanide. Is NADPH produced in this process? Explain.

**23.** *How Often Does a Chlorophyll Molecule Absorb a Photon?* The amount of chlorophyll *a* ($M_r$ 892) in a spinach leaf is about 20 $\mu$g/cm$^2$ of leaf. In noonday sunlight (average energy 5.4 J/cm$^2$ • min), the leaf absorbs about 50% of the radiation. How often does a single chlorophyll molecule absorb a photon? If the average lifetime of an excited chlorophyll molecule in vivo is 1 ns, what fraction of chlorophyll molecules are excited at any one time?

**24.** *Effect of Monochromatic Light on Electron Flow* The extent to which an electron carrier is oxidized or reduced during photosynthetic electron transfer can sometimes be observed directly with a spectrophotometer. When chloroplasts are illuminated with 700 nm light, cytochrome *f*, plastocyanin, and plastoquinone are oxidized. When chloroplasts are illuminated with 680 nm light, however, these electron carriers are reduced. Explain.

**25.** *Function of Cyclic Photophosphorylation* When the [NADPH]/[$NADP^+$] ratio in chloroplasts is high, photophosphorylation is predominantly cyclic (Fig. 18–44). Is $O_2$ evolved during cyclic photophosphorylation? Explain. Can the chloroplast produce NADPH this way? What is the main function of cyclic photophosphorylation?