nm. The internal structure includes a helical arrangement of nucleoprotein. They also are negative strand viruses.

Among the **retroviruses**[512–514] are types B and C **oncoviruses** which induce malignant tumors in mammals and birds[515] and the **human immunodeficiency virus** (HIV), the apparent causative agent of AIDS. Their RNA functions in a surprising way. Each virion contains a **reverse transcriptase**, an enzyme that transcribes copies of circular dsDNA copies from the one or two mRNA-like molecules that make up the virus genome. Following action of the reverse transcriptase, one of the transcribed DNA circles becomes covalently spliced into the host's own cellular DNA. There it remains permanently as a **provirus**. RNA molecules transcribed from the provirus serve as mRNA for virus-encoded proteins and also as the genomes for new virus particles.

Double-stranded RNA is unusual in nature but constitutes the genome of the **reoviruses**.[516] The RNA of these viruses fragments into segments upon infection. One member of the group is thought to be the cause of acute diarrhea of infants.[517]

## 4. Viruses without Nucleic Acid?

The cause of the slow, fatal neurological disease of sheep known as **scrapie** has been a mystery for many years. Similar human diseases include **kuru** and **Creutzfeldt-Jakob** disease.[518–520] Scrapie can be transmitted by injection and this has permitted isolation of the apparent infective agent, a 27- to 30-kDa hydrophobic protein particle[521] which is devoid of DNA or RNA. Prusiner[521] suggested the name **prion** (proteinaceous infectious particle) for the scrapie agent. However, mRNA for the prion is present in normal as well as infected brains, and protein produced in mouse cells from cloned prion genes did not cause scrapie infections. Therefore, there was doubt about the causative agent for the disease. The prion concept is now generally accepted and is considered further in Chapter 29. There are still some who are looking for a nucleic acid component.[519,522–524]

## 5. Life Cycles

Viruses have many modes of life. They enter cells in various ways. Some enter through coated pits from which they are taken into lysosomes via endocytosis. Others are literally injected into the cells (See Box 7-C). Within cells some viruses are assembled in the nucleus, some in the cytoplasm, and some in membranes. The typical life cycle of a virus leads to rapid formation of large numbers of progeny. Within 20 minutes after entrance into a bacterial cell, a bacteriophage can induce the formation of 100–200 new bacteriophage particles. One of the bacteriophage genes encodes a protein that is also synthesized by the host and which induces lysis of the cell membrane and destruction of the cell. Many animal viruses destroy cells in a similar fashion.

**Temperate bacteriophage**, the best known being phage λ, have a very different life cycle. Their DNA usually becomes integrated at a specific point into the genome of the bacterium (Chapter 27). Only rarely is an infected cell lysed. The retroviruses that attack mammals and birds have a similar characteristic. Their DNA is also integrated into the host genome. Some viruses that usually produce lysis of cells, e.g., SV40, adenoviruses, herpes viruses, and hepatitis B virus, can occasionally be integrated into the DNA of the host. If such integration occurs in the middle of a gene, that gene will be mutated. This is one way in which such viruses may induce cancers.

One of the most important results of integration of viral DNA into the host genome is that the integrated genes are replicated as part of the genome and are transmitted from one generation to the next. Among these are the cancer-causing **viral oncogenes** (*v-onc*), which are discussed in Box 11-D and in Chapter 11, Section H. While viruses are important causes of cancer in some animals, relatively few human cancers are thought to result directly from the action of viruses. However, the Epstein–Barr virus, which causes mononucleosis, can sometimes be integrated into epithelial cells of nasal regions and can evidently cause cancer. The same virus appears to be responsible for Burkitt's lymphoma, a common cancer in certain areas in Africa.[525]

## 6. Plasmids and Transposable Genetic Elements

In addition to their chromosomal DNA, bacteria often carry extra small pieces of DNA as permanent parts of their genome. These **plasmids** (sometimes called **episomes**), which are about the size of the DNA of viruses, replicate independently of the host chromosomes. Each bacterial cell usually contains more than a single copy of the plasmid. For example, the "colicinogenic" plasmid **ColE1**, that infects *E. coli* is a circular piece of DNA of molecular mass $4.2 \times 10^6$ Da. Over 20 copies are normally found per cell but in the presence of a suitable concentration of the drug chloramphenicol the number may rise to 1000–2000.

Plasmids carry a variety of genes which are often useful to bacteria. Some proteins encoded by plasmid genes confer drug resistance to a bacterium. Some are antibiotics. For example, a protein encoded by a gene in plasmid ColE1 is toxic to other strains of *E. coli.* Some plasmids carry genes for enzymes needed for the oxidation of hydrocarbons. Some plasmids contain

genes for the **restriction endonucleases** which have become essential to present-day molecular biology and genetic engineering (Section H, 2).

As with some viruses, the DNA of many plasmids can become integrated into the genome of the host. An example is provided by the large 62-kDa plasmids known as **sex factors**. They contain genes encoding the protein subunits of the sex pili (Chapter 7) and can become integrated into the bacterial chromosome. Bacteria containing integrated sex factors are "male" and are able to transfer genes not only of the sex factor but also of virtually the entire bacterial genome into other susceptible bacterial cells. This provides bacteria with the means for sexual reproduction. The transfer of DNA between the bacteria may occur via the sex pili (see Chapter 26). In this respect the sex factors are similar to viruses such as M13 that also appear to gain entrance to bacteria via sex pili.[72]

Integrated viruses are also related to **transposable genetic elements** (transposons). These are segments of DNA that allow genes to move from place to place within the chromosomes (Chapter 27).

## H. Methods of Study

Many of the methods discussed in Chapter 3 are directly applicable to nucleic acids. A few additional methods will be considered in this section.[525a]

### 1. Isolation and Separation of Nucleic Acids

RNA is often extracted from lysed cells or tissues, separated ribosomes, mitochondria, plastids, or nuclei by warming with aqueous phenol and a detergent such as sodium dodecyl sulfate (SDS). Proteins are denatured by this treatment and are dissolved by the phenol, while RNA remains in the lighter aqueous layer. Depending on the conditions DNA may either remain in the aqueous layer or be removed.[526–528] Various precipitation and extraction procedures may be used to separate the RNA in the aqueous layer from polysaccharides, from DNA (if present), and from their components.[526,528,529] DNA may be extracted from cells or nuclei as a nucleic acid–protein complex using 1 M NaCl. The protein can then be denatured with an organic solvent, by detergents, or by phenol. It is desirable to digest proteins away with a nonspecific protein-hydrolyzing enzyme such as proteinase K.[528] After removal of proteins DNA is often precipitated with cold ethanol.

During isolation of RNA, bentonite (a type of clay) or other inhibitors of ribonuclease are often added. For the same reason, chelating agents that complex metal ions needed for the action of deoxyribonucleases are used to protect DNA. Care is necessary to avoid shearing of the very long, narrow strands of DNA. Even rapid pipetting of solutions will cause such breakage.

Extracted RNA molecules may be separated from each other by centrifugation in a sucrose gradient (Chapter 3).[528] Fragments of DNA are purified in the same way or by equilibrium centrifugation in **CsCl gradients**.[37] Concentration gradients in the dense salt solution are stable, and the sharpness of banding of particles is ensured by use of a high centrifugal field. Single-stranded DNA may be separated from double-stranded DNA, and DNAs of differing G + C content can be separated. The latter separation is based on differences in buoyant densities ρ in CsCl which are approximately shown in Eq. 5-9.

$$\rho = 1.660 + 0.098 \text{ (mole fraction C + G)} \qquad (5\text{-}9)$$

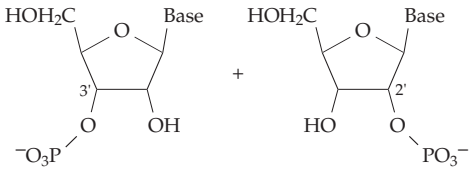
One of the most important methods for separating either RNA or DNA mixtures is zone electrophoresis through polyacrylamide or agarose gels. The separated bands may be visualized by scanning in ultraviolet light or by fluorescence of intercalated dyes such as ethidium bromide (Figs. 5-20, 5-22). This method is being displaced to some extent by HPLC using DEAE type ion exchange columns[530,531] for small lengths of DNA including plasmids. The procedure called **pulsed field electrophoresis** makes it possible to isolate very large pieces of DNA, up to several million base pairs in length.[532–534] The separation is carried out in agarose gels, through which the long DNA rods must move in a snakelike fashion.[534] The current is delivered in a pulse and then, after a period of a second to several minutes, a second pulse in a different direction, usually at 90° to the first. The procedure is repeated many times. The size of the DNA seems to affect the time required to reorient the molecules and to start moving in the second direction. Intact DNA from small chromosomes can be separated (Fig. 5-42). To prevent breakage of the DNA by shearing, intact cells are suspended in liquid agarose and allowed to gel into a block about 2 x 5 x 10 mm in size. The block is treated with enzymes and detergents to lyse the cells and to remove all protein and RNA.[534] The block, containing the residual DNA molecules, is then embedded in the electrophoresis gel. Other methods of DNA separation include chromatography on hydroxyl-apatite and gel filtration.

### 2. Hydrolysis and Analysis

Both DNA and RNA are easily broken down by acid-catalyzed hydrolysis. Thus, heating at 100°C for one hour in 12 M $HClO_4$ is sufficient to hydrolyze nucleic acids to their constituent bases. However, for analysis of RNA it is better to heat in 1 N HCl for 1 h at

100°C. The products are adenine, guanine, cytidine-5'-phosphate, and uridine-5'-phosphate.[535] As is suggested by this distribution of products, the glycosylamine linkages to purines are more labile than those to pyrimidines. The linkages are also less stable in DNA than in RNA. A procedure based on these differences and useful in sequencing by the Maxam-Gilbert method, is to leave DNA overnight in the cold at pH 2 to cleave off all of the purine bases. The resulting polymer is known as an **apurinic** acid.

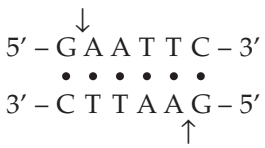
In alkaline solutions RNA is hydrolyzed to a mixture of 2'- and 3'-nucleotides.




The mechanism involves participation of the free 2'-OH of the ribose groups and formation of cyclic 2', 3'-phosphates and is similar to that of pancreatic ribonuclease (Chapter 12). Because deoxyribose lacks the free 2'-OH, the phosphodiester linkages in DNA are quite stable in base.

Hydrolytic cleavages of nucleic acids by the enzymes known as **nucleases** are of great practical value. Pancreatic ribonuclease, an **endonuclease**, cuts a chain adjacent to a pyrimidine in nearly random fashion, leaving phospho groups attached to the 3' position in the nucleotide products (Fig. 5-43). **Exonucleases** cleave from the ends of chains. For example, the phosphodiesterase of snake venom cleaves from the 3' end, which must have a free 3'-OH group, to give 5'-nucleotides. On the other hand, the phosphodiesterase from spleen has the opposite polarity, cleaving chains from the 5' end to give 3'-nucleotides. Similar variations in specificity are found among enzymes that cleave DNA. For example, pancreatic DNAase I, which cleaves preferentially between adjacent purines and pyrimidines, yields 5' mononucleotides whereas DNAase II gives 3'-mononucleotides. Various hydrolytic cleavage reactions of polynucleotides are summarized in Fig. 5-43.

The most striking specificity in DNA hydrolysis is displayed by the **restriction endonucleases** which are discussed further in Chapter 26. These fussy catalysts cleave only at points within or close to a defined sequence of several nucleotides in double-stranded DNA. For example, the enzyme EcoR I cuts only at the following palindromic sequence:



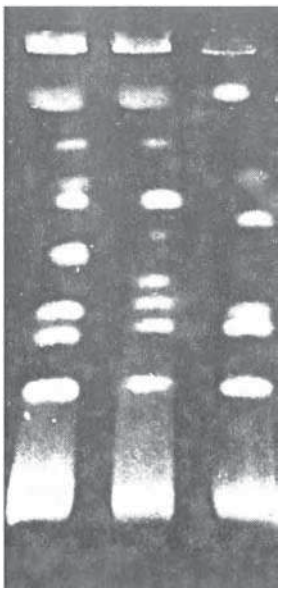
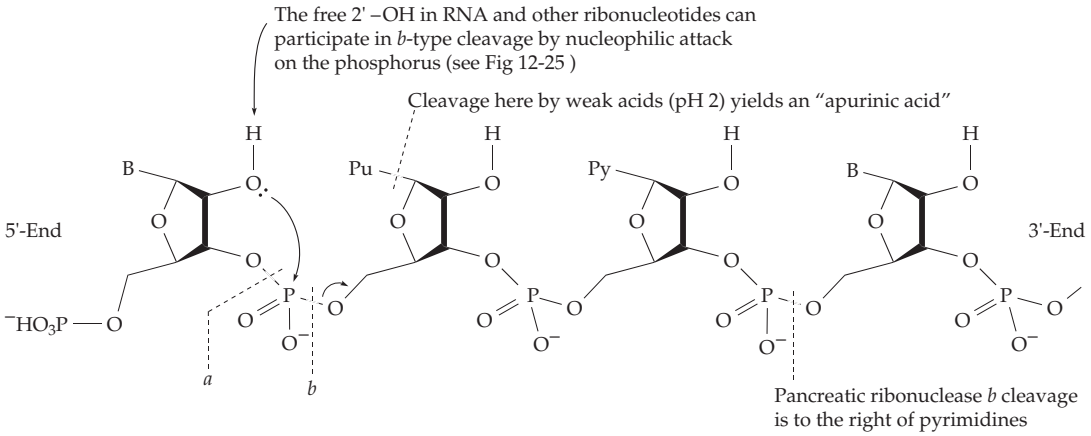



**Figure 5-42** Intact DNA from the chromosomes of three strains of the malaria parasite *Plasmodium falciparum*, ranging from 750 Kb to 5 Mb, separated by pulsed-field gel electrophoresis. Courtesy of C. Smith and T. E. Wellems. Reproduced by permission of Amersham Pharmacia Biotech Inc.

The cuts in the two strands are made at the points indicated by the arrows. This one endonuclease will cut almost any DNA into long pieces averaging about 5000 base pairs each. These pieces can in turn be cleaved by other restriction endonucleases to form smaller fragments. Since there are about 2400 of these enzymes known, with 188 different specificities,[536] it is possible to cut any piece of DNA down to a size of 100–500 base pairs, ideal for sequencing.[537–539] Each fragment has known sequences at the two ends. Some restriction enzymes cleave outside their specific recognition sequence (see Table 26-2). Some recognize 16-nucleotide palindromes and cut at rare sites.

It is sometimes desirable to cut a large DNA molecule at only a few points. One approach is to protect most sites of a restriction enzyme's action by methylating them (see Chapter 26) while protecting the desired cleavage site, for example by a repressor protein[540] or by a PNA molecule (p. 227) of specific sequence designed to "clamp" the site chosen for protection.[541] Ribozymes (Chapter 12) have been engineered to be as specific or more specific than endonucleases.[542,543] Other new approaches are being developed.[544]

The base composition of either RNA or DNA can be determined after hydrolysis catalyzed by 98% formic acid at 175°C for 30 min or by 12 M perchloric acid at 100°C for 1 h.[545] The bases can then be separated by ion exchange chromatography on a sulfonated polystyrene resin. RNA can be hydrolyzed to a mixture of nucleoside 2'- and 3'-phosphates by 0.3 M NaOH at




A. Cleavage at point *a* is catalyzed
   1. Throughout the molecule by endonucleases
      Pancreatic deoxyribonuclease I
   2. Only at the 3' end by exonucleases
      Venom diesterase, nonspecific, attacks DNA and RNA. A free 3'-OH is essential

B. Cleavage at point *b* is catalyzed
   1. Randomly throughout the molecule by endonucleases and by bases (nonenzymatically)
      Pancreatic ribonuclease cleaves only to the right of a pyrimidine-containing nucleotide
      Ribonuclease T1 of *Aspergillus oryzae* cleaves to the right of a guanine-containing residue (3'-guanylate) in ssRNA
      Ribonuclease T2 of *Aspergillus oryzae* cleaves to the right of an adenine-containing residue (3'-adenylate) in ssRNA
      Pancreatic deoxyribonuclease (DNase) II
      Micrococcal DNase
   2. Only at the 5' end by exonucleases
      Bovine spleen phosphodiesterase hydrolyzes both polyribo- and polydeoxyribonucleotides

**Figure 5-43** Some hydrolytic cleavage reactions of polynucleotides. Reactions of both RNA and DNA are included.

37°C for 16 h and DNA can be hydrolyzed to nucleotides enzymatically. The negatively charged nucleotides can then be separated by ion exchange chromatography on a quaternary base-type resin (Chapter 3). Periodate cleavage (Eq. 4-12) and reduction of the resulting dialdehydes by [$^3$H]$NaBH_4$ to trialcohols allows introduction of a radioactive label (Fig. 5-44). Alternatively, the dialdehydes can be reductively alkylated by an amine plus $NaCNBH_3$.[546]

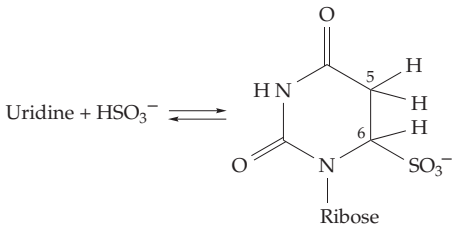
The total content of RNA + DNA in tissues may be estimated from the phosphorus content or by color reactions of the sugars.[37,545] These reactions depend upon dehydration to furfural or deoxyfurfural by concentrated sulfuric acid or HCl (Eq. 4-4). Furfural formed from RNA reacts with orcinol (3,5-dihydroxytoluene) and ferric chloride to produce a green color useful in colorimetric estimation of RNA. A similar reaction of DNA with diphenylamine yields a blue color.

Quantitative determination of over 90 free nucleotide compounds found within cells can be accomplished by thin layer chromatographic procedures on as few as $10^6$ bacterial cells (~2 μg) labeled by growth in a $^{32}P_i$-containing medium.[547]

## 3. Characteristic Reactions of the Bases and Backbone

***Reactions of nucleophiles.*** A number of nucleophilic reagents add reversibly at the 6 position of pyrimidines. Thus, bisulfite adds to uridine (Eq. 5-10).[528] Hydroxylamine ($HONH_2$) adds in a similar fashion to give a compound with –HNOH in the 6 positions.[528] Sodium borohydride ($NaBH_4$), which can be viewed as a donor of a hydride ion ($H^-$), reduces uridine to the 5,6-dihydro derivative. This presumably occurs by attack of the hydride ion at position 6 in a manner analogous to the reaction of bisulfite in Eq. 5-10.

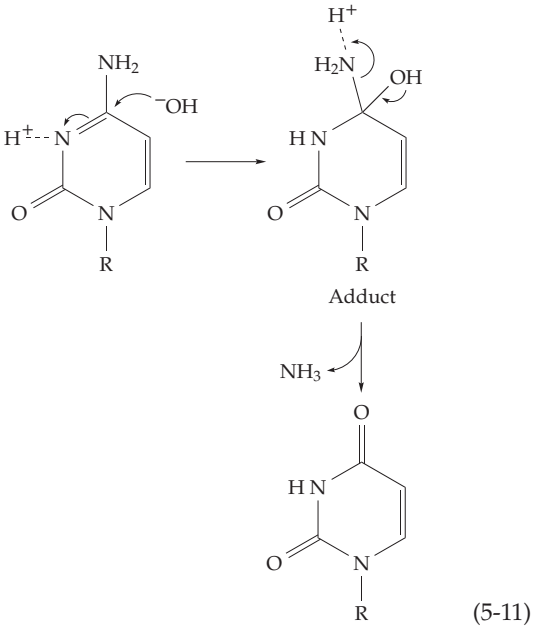
Cytidine reacts in the same way, but the bisulfite addition compound is unstable. These C5–C6 adducts of cytidine all have a greatly enhanced reactivity at C4,



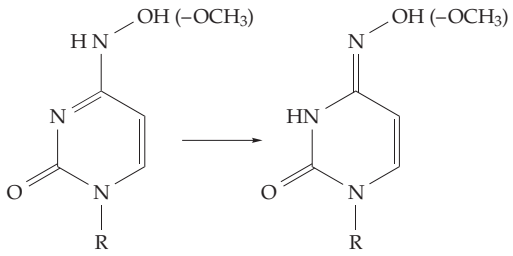

(5-10)

presumably because of the lessened aromatic character of the ring. Cytidine is slowly deaminated by base, presumably as a result of attack by hydroxyl ion on the electrophilic center at C4 and subsequent elimination of $NH_3$ (Eq. 5-11). The reaction is catalyzed by buffer salts and by bisulfite and hydroxylamine. Catalysis probably occurs, at least in part, as a result of addition of these nucleophiles at the 6 position to form compounds with increased nucleophilic reactivity at C4.[548] Hydroxylamine and methoxyamine ($NH_2OCH_3$) participate in reactions parallel to that of the hydroxyl ion in Eq. 5-11. Products contain –NHOH or $-NH-OCH_3$ in place of $-NH_2$ but tautomerize to the more stable forms shown in Eq. 5-12. Similar substitution reactions occur with other amines.[33] The C6 adduct with hydrazine can undergo ring cleavage (Eq. 5-13). The initial product then undergoes β elimination, leaving ribosylurea or deoxyribosylurea. The same reaction can be carried out on intact strands of DNA and is widely used in determination of nucleotide sequences.

The conversion of 5-hydroxymethylcytosine to the




(5-11)




(5-12)

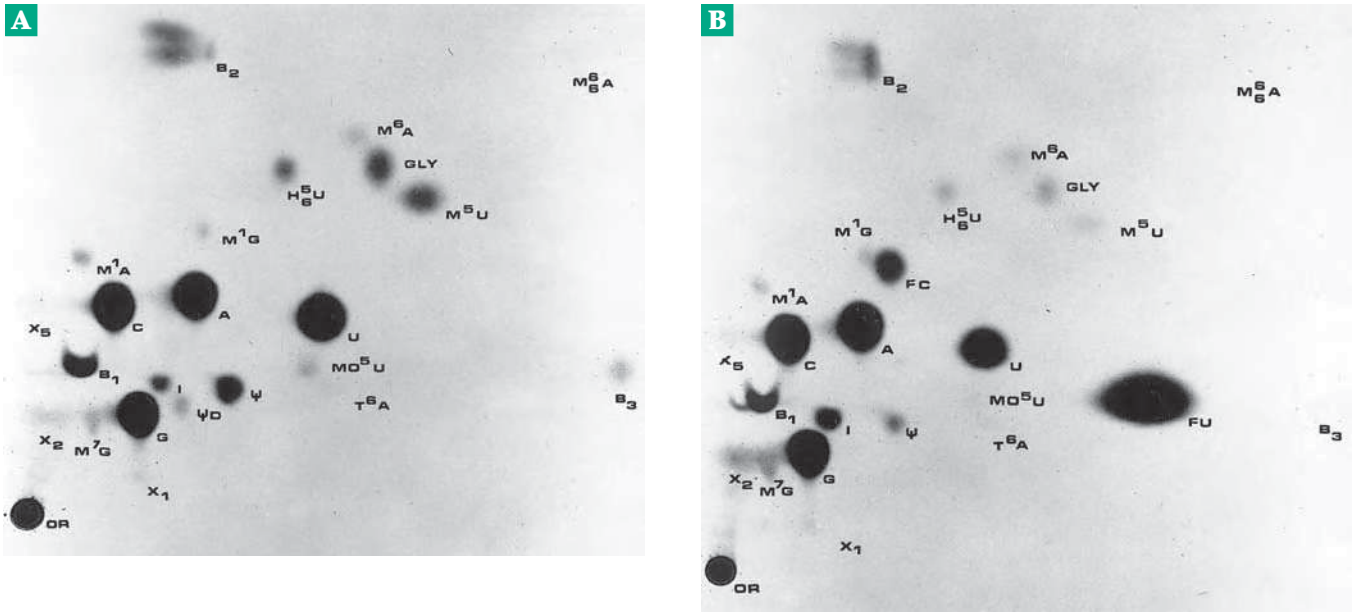
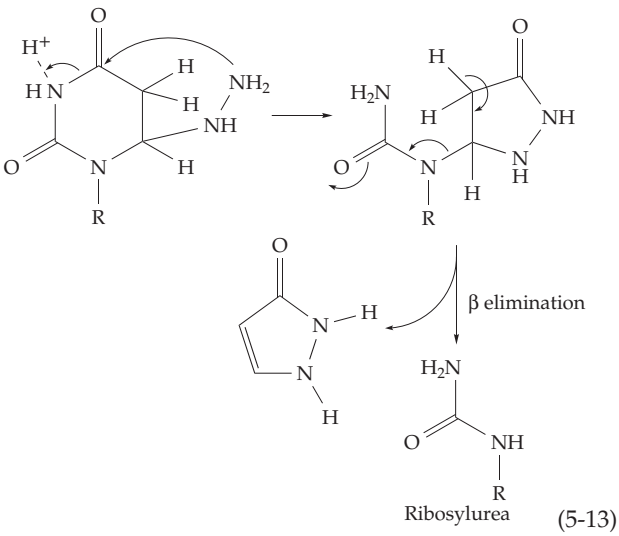



**Figure 5-44** Fluorographs of 2′3′-[$^3$H] nucleoside trialcohols from *Bacillus subtilis* grown in the absence (A) and presence (B) of 5-fluorouracil. About 2.3 nmol of nucleosides from each sample was chromatographed and exposed to X-ray film for 90 h. at –80 °C. Or (origin) $B_1$, $B_2$, and $B_3$ contain unidentified material present in a reaction mixture lacking RNA. Abbreviations used: FU, 5-fluorouridine; FC, 5-fluorocytidine; U, uridine; C, cytidine; G, guanosine; A, adenosine; I, inosine; m′G, 1-methylguanosine; $m^7G$, 7-methylguanosine; m′A, 1-methyladenosine; $m^6A$, 6-methyladenosine; $m^6{}_6A$, 6,6-dimethyladenosine; $t^6A$, N-[9-(β-D-ribofuranosyl) purin-6-yl carbamoyl] threonine; $H^5{}_6U$, 5,6-dihydrouridine; ψ, pseudouridine; $\psi_D$, decomposition product of ψ; $m^5U$, 5-methyluridine (ribosylthymine); $mo^5U$, 5-methoxyuridine; N′, a nucleoside trialcohol obtained by reduction of a nucleoside dialdehyde with [$^3$H]$NaBH_4$; FU-5 and FU-20 samples correspond to tRNAs from cells grown at that final concentration of 5-fluorouracil in μg / ml. Courtesy of Ivan Kaiser.

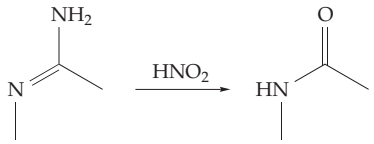



(5-13)

5-methylenesulfonate (5-$CH_2SO_3^-$) by reaction with bisulfite should also be mentioned.[549] This is a nucleophilic displacement on the electron-deficient methylene group of this base.

***Reactions with electrophilic reagents.*** Reactions of nucleic acids with the simplest electrophile, the proton, have been considered in Section A2. Somewhat similar are the reactions by which metal ions bind at many sites on both the bases and the phosphate groups of the backbone.[550]

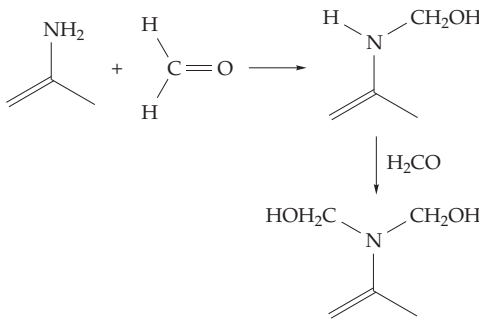
An important reaction is the deamination of amines by dilute **nitrous acid**. This reagent, by a complex mechanism, converts the amino groups of cytidine, adenosine, and guanosine to hydroxyl groups; hydroxy compounds tautomerize to the corresponding amides (Eq. 5-14).




(5-14)

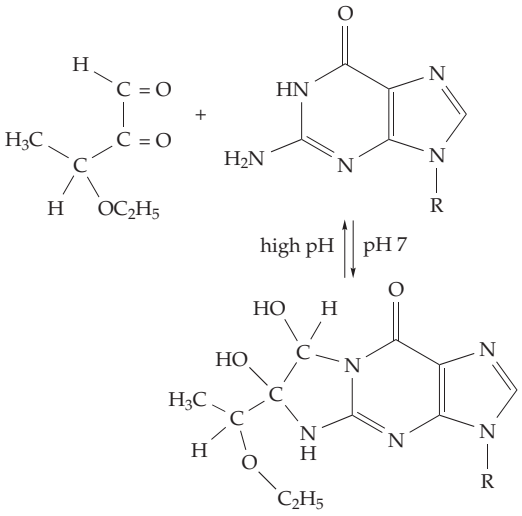
Cytidine reacts more rapidly than does adenosine which in turn reacts more rapidly than guanosine. The reaction converts cytosine into uracil and adenine into hypoxanthine. The changes are mutagenic because during replication the modified bases of the DNA pair differently than do the original bases. Guanine is converted to xanthine but this is not likely to be highly mutagenic. Nitrous acid can also convert uridine to 5-nitrouridine.

The amino groups of the bases react reversibly with aldehydes but to a lesser extent than do the more strongly basic amino groups of the amino acids. Formaldehyde forms adducts containing either one or two molecules of the aldehyde (Eq. 5-15).




(5-15)

These are reversible reactions. A more nearly irreversible crosslinking can occur by elimination of water between one of these products and a nucleophilic group in another base. A dicarbonyl reagent that is widely used because of its specificity toward guanine is **kethoxal** (Eq. 5-16).




(5-16)

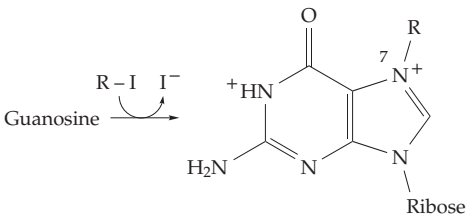
Formation of the cyclic product is a consequence of the presence of the adjacent amino and NH groups in the guanine ring.

Pyrimidines undergo **halogenation** at position 5 (Eq. 5-17), while guanine reacts at position 8. Adenine is quite unreactive. Elemental halogens or a variety of other halogenating reagents may be used. Of special

$$Br_2 + \text{cytidine} \rightarrow \text{5-Br-cytidine} + H^+ + Br^- \quad (5\text{-}17)$$

value is iodination with $^{131}I$ or $^{125}I$, by which a high level of radioactivity may be introduced into nucleic acids.

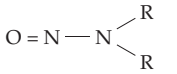
**Alkylation** reactions are not only of use in structural studies but also provide the basis for the action of a large class of mutagenic compounds.[528] Treatment of a nucleoside, nucleotide, or nucleic acid with an alkyl iodide or a dialkylsulfate converts residues of guanosine to an $N^7$-alkyl-guanosine (Eq. 5-18).




(5-18)

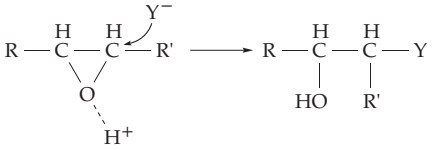
Reaction occurs at other nitrogen atoms as well as the oxygen atom of the base and of the ribose ring to a lesser extent. Adenine is alkylated preferentially at N-1 and cytosine and thymine at the corresponding position (N-3) almost exclusively. Uridine and thymidine react very slowly. Adenine is also alkylated at N-3, N-7 and at the exocyclic N-6.

Other alkylating reagents include the powerful mutagens dialkylnitrosamines and alkylnitrosoureas.



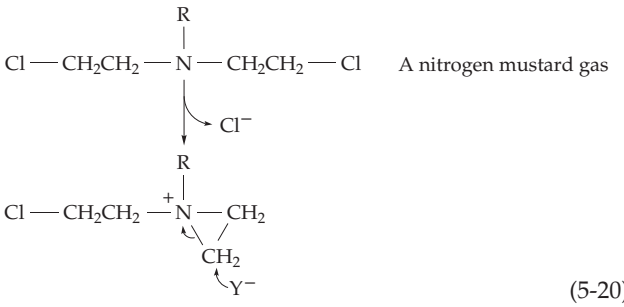

Dialkylnitrosamines

Epoxides alkylate by a nucleophilic displacement reaction that opens the ring (Eq. 5-19).



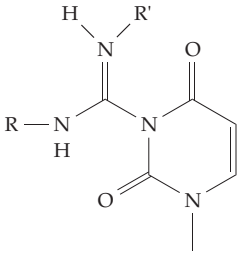

(5-19)

The nitrogen and sulfur mustards undergo internal ring closure to an iminium ion (Eq. 5-20) which can then open by attack of a nucleophilic atom of the nucleic acid.



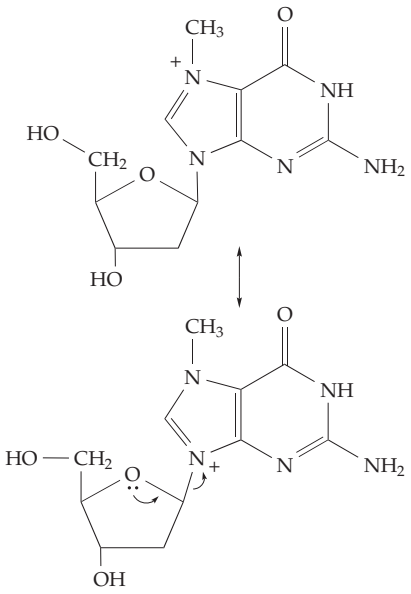

(5-20)

Other alkylating agents react through nucleophilic addition to a carbon–carbon double bond. Thus, acrylonitrile reacts with the nitrogen or oxygen atoms of nucleic acids in the same manner as does the SH group in Eq. 3-25. The water-soluble carbodiimides react as in the first step of Eq. 3-10 to form adducts of the following type:



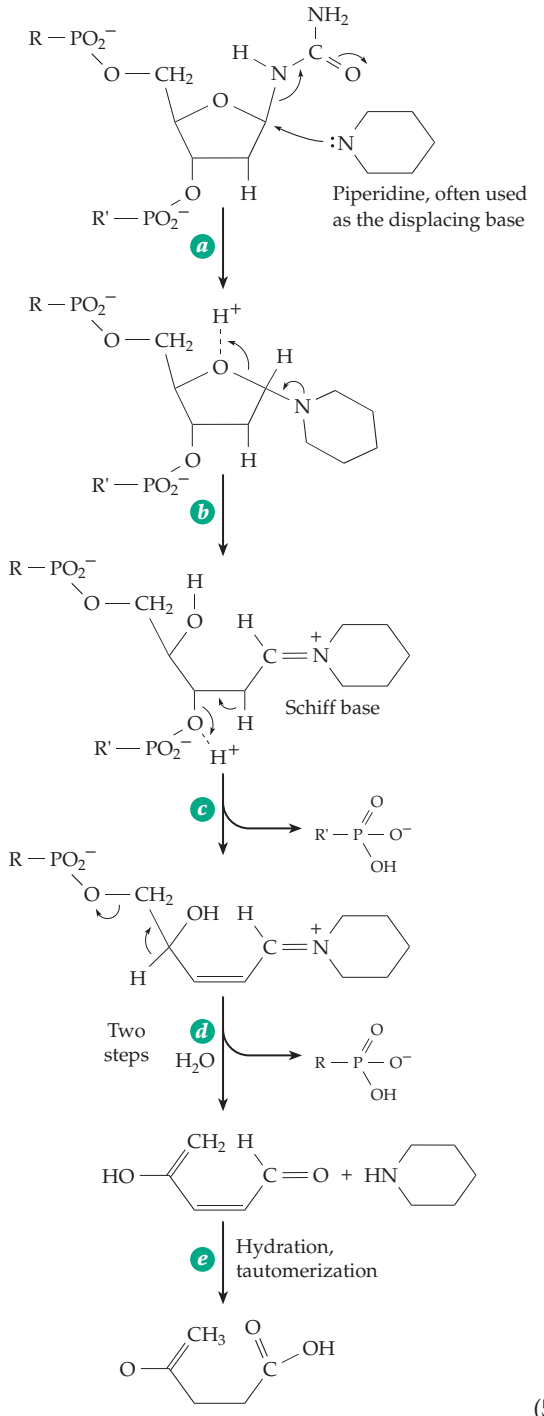

There are many other alkylating agents which often display widely varying reactivity and specificity toward particular nucleic acid bases and particular nucleotide sequences.

A striking effect of alkylation of guanine in nucleic acids is the labilization of the *N*-glycosyl linkage to the ribose or deoxyribose. This effect can be understood in terms of the induction by resonance of a partial positive charge on the nitrogen of the glycosyl linkage.




As is indicated by the small arrows on the right-hand structure, the positive charge assists in an elimination reaction that produces an oxycarbocation. The latter can then react with a hydroxyl ion from water.

***Reactions causing cleavage of the sugar-phosphate backbone.*** Treatment of DNA with 16–18 M hydrazine (Eq. 5-13) leads to the destruction of the pyrimidine rings. The reaction can be made somewhat specific for cytosine by carrying it out in the presence of a high concentration of chloride.[551] The remaining polymer, an apyrimidinic acid, contains residues of ribosylurea. These undergo an amino-catalyzed displacement and a β elimination sequence that cleaves the polynucleotide chain (Eq. 5-21). Hydration of the aldehyde (Eq. 13-1) and several




(5-21)

tautomerization steps are involved in step *e* of this equation. This reaction is very useful in sequence determination (Section 6). Notice that "tracts" of purine nucleotides remain intact after this treatment. A similar base-catalyzed reaction sequence can be used to displace $N^7$-methylguanine and to cleave the polynucleotide. Ethylnitrosourea, in its reaction with purines, is useful as a structural probe of RNA.

## 4. Melting, Hybridization, and Polynucleotide Probes

Like proteins, nucleic acids can undergo denaturation. The strands of the double helix of DNA are separated and the double-stranded regions of RNA molecules "melt." Denaturation can be accomplished by addition of acids, bases, and alcohols or by removal of stabilizing counter ions such as $Mg^{2+}$. The product is a random coil and denaturation can be described as a helix → coil transition. Denaturation of nucleic acids by heat, like that of proteins, is cooperative (Chapter 7, Section A,3) and can be described by a characteristic **melting temperature**.

A plot of the optical absorbance at 260 nm (the wavelength of maximum light absorption by nucleic acids) versus temperature is known as a **melting curve** (Fig. 5-45). The absorbance is lower, by up to 40%, for native than for denatured nucleic acids. This **hypochromic effect** (Chapter 23) is a result of the interaction between the closely stacked bases in the helices of the native molecules. The melting temperature $T_m$ is taken as the midpoint of the increase in absorbance (Fig. 5-45). As the percentage of G + C increases, the nucleic acid becomes more stable toward denaturation because of the three hydrogen bonds in each GC pair. $T_m$ increases almost linearly with increases in the G + C content. In the "standard" citrate buffer (0.15 M NaCl + 0.015 M sodium citrate, pH 7.0) Eq. 5-22 holds. The exact numerical relationship depends strongly upon the ionic composition and pH of the medium.[37,72,552,553]

$$\%\,(G + C) = 2.44\,(T_m - 69.3);\ T_m \text{ in } °C \qquad (5\text{-}22)$$

The curves in Fig. 5-45 appear simple, but using newer apparatus and plotting the first derivative of the melting curve yields a complex pattern that depends on the sequence of bases.[555]

Complete denaturation of DNA leads to separation of the two complementary strands. If a solution of denatured DNA is cooled quickly, the denatured strands remain separated. However, if the temperature is held for some time just below $T_m$ (a process known as **annealing**), the native double-stranded structure can be reformed. An important tool for studying DNA has been the measurement of the **kinetics of reassociation** of separated strands of relatively short DNA fragments.[72,556,557]

Because it depends upon the concentration of two separated strands, reassociation obeys second-order kinetics (Chapter 9) and Eq. 5-23, which is readily derived by integrating Eq. 9-8 for [A] = [B] = C from time 0 to *t*:

$$C / C_0 = 1 / (1 + k\,C_0 t) \qquad (5\text{-}23)$$

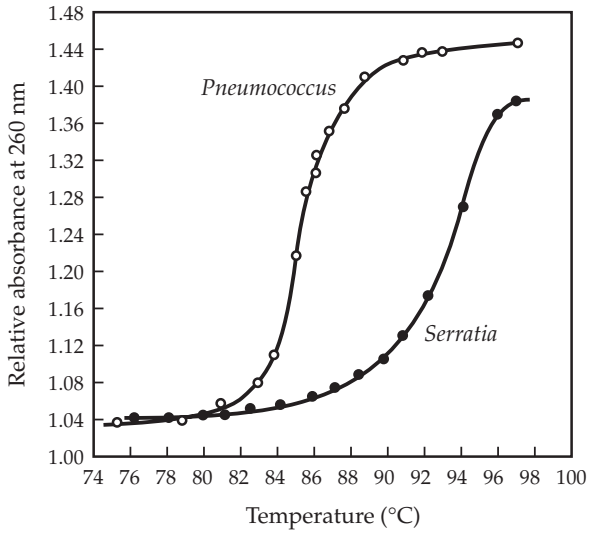



**Figure 5-45** A melting curve for DNA molecules from two different sources. From Davidson.[554]

The initial concentration of denatured DNA, $C_0$, is related in this way to the concentration C of DNA remaining dissociated at time $t$. A plot of the fraction of molecules remaining single-stranded versus the logarithm of $C_0t$ (Fig. 5-46A) is a convenient way of displaying data. As indicated in Fig. 5-46B, the value of $C_0t$ increases in direct proportion to the length of the DNA chain in the genome, but it is very much decreased if the sequence of bases is highly repetitive [poly(T) and poly(A)]. The slope of the plot at the midpoint gives an indication of the heterogeneity of the DNA fragments in a solution.

Denatured DNA fragments can sometimes reassociate with DNA from a different source to form **hybrid duplexes**. Such double helices, in which one strand comes from one strain of an organism and the other strand from a genetic variant of the same organism or from a different species, are known as **heteroduplex**. Some mutations consist of **deletions** or **additions** of one or a substantial number of bases to a DNA chain. Heteroduplexes prepared from DNA of such mutants hybridized with that from a nonmutant strain have normal hydrogen-bonded Watson–Crick base pairs for the most part. However, they may have single-stranded loops in regions where long deletions or additions prevent complementary base pairing.

Hybridization measurements have been used in many studies of **homology** of nucleic acids from different species. A nucleic acid is cut (e.g., by sonic oscillation) into pieces of moderate length (~ 1000 nucleotides) and is denatured. The denatured DNA fragments are mixed with denatured DNA of another species. Nucleotide sequences that are closely similar between species tend to hybridize, whereas sequences that are drastically different between two species do not (Fig. 5-46). One way to do such an experiment is to immobilize the long-chain denatured DNA from the one organism by embedding it in an agar gel[558] or by absorbing it onto a nitrocellulose filter.[559,560] The DNA fragments from the second organism are passed through a column containing "beads" of the DNA-containing agar or through the filter with adsorbed DNA. Pairing of fragments with complementary sequences occurs

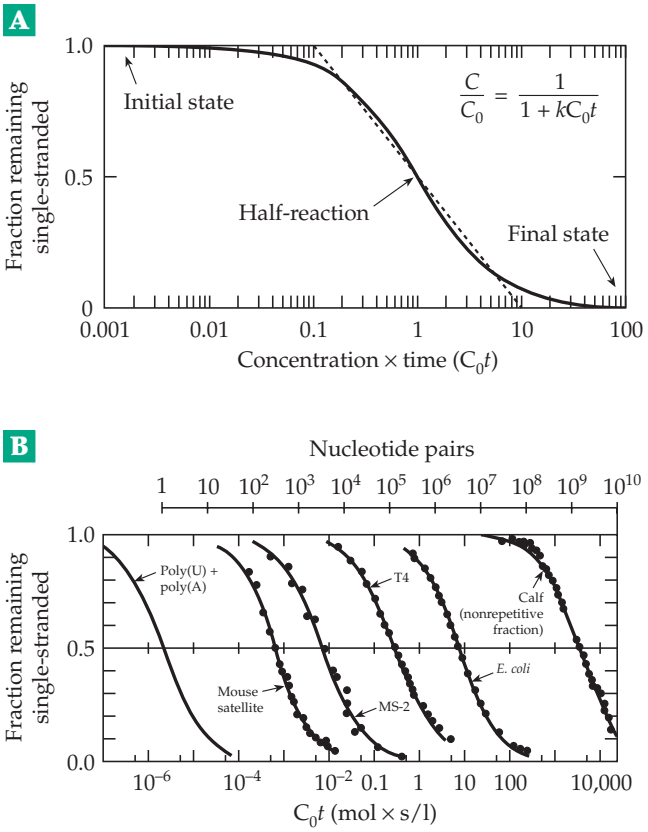


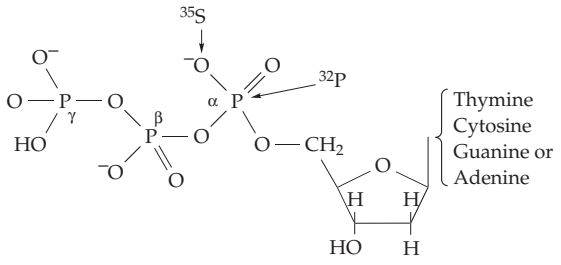

**Figure 5-46** Reassociation curves for DNA from Britten and Kohne.[556,561] (A) Time course of an ideal, second-order reaction to illustrate the features of the log $C_0t$ plot. The equation represents the fraction of DNA which remains single-stranded at any time after the initiation of the reaction. For this example, $k$ is taken to be 1.0, and the fraction remaining single-stranded is plotted against the product of total concentration and time on a logarithmic scale. (B) Reassociation of double-stranded nucleic acids from various sources. The genome size is indicated by the arrows near the upper nomographic scale. Over a factor of $10^9$, this value is proportional to the $C_0t$ required for half-reaction. The DNA was sheared, and the other nucleic acids are reported to have approximately the same fragment size (about 400 nucleotides, single-stranded). Correction has been made to give the rate that would be observed at 0.18 M sodium-ion concentration. The temperature in each case was optimal, i.e., ~30°C below the melting temperature $T_m$. The extent of reassociation was established by measuring optical rotation (calf thymus DNA), ribonuclease resistance (MS-2), or hypochromicity.

and such paired fragments are retained while strands that do not pair pass on through the column (or filter).

Both DNA hybrids and **DNA–RNA hybrid duplexes** are very important to present day genetic research.[560,562] Molecules of mRNA that represent transcripts of a particular gene will hybridize only with one of the two separated strands of DNA for that gene.

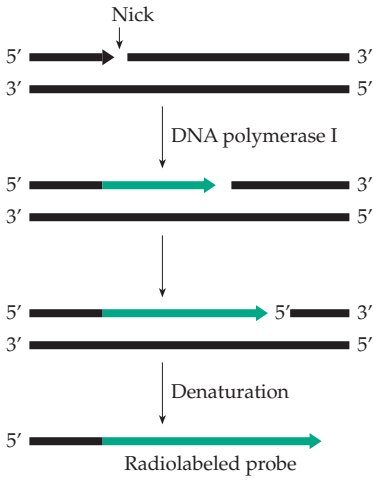
A major use of hybridization is to locate a gene or other DNA or RNA sequence by means of a **synthetic probe**.[563] This is a small piece of DNA or RNA which is labeled in some way, e.g., with a radioisotope such as $^{3}H$, $^{32}P$, or $^{125}I$. Alternatively, the probe may carry a highly fluorescent dye or a "tag" that can be recognized by a specific antibody.[564,565] An example of the latter is the use of the vitamin biotin and the specific binding protein **avidin** (see Box 14-B).[566] Related procedures employ labeling with **digoxigenin** and often employ chemiluminescent detection.[567–569] Several methods for preparation of probes are in use. Some are enzymatic but the direct chemical synthesis of oligonucleotide probes is probably used the most.

One of the first methods devised for making a highly radioactive DNA probe is called **nick translation**.[570] A piece of dsDNA, e.g., a "restriction" fragment cut from a larger piece of DNA by restriction endonucleases, is selected. A small amount of pancreatic DNase I is added. It creates "nicks" in which one strand has been cut and some nucleotides have been removed leaving a gap. Now the DNA is incubated with DNA polymerase I (pol I) and a mixture of the four mononucleotide triphosphates, the precursors of biological synthesis of DNA (Chapter 27). Usually a $^{32}P$ or $^{35}S$ label with high radioactivity is present in one of the nucleotide triphosphates as indicated in the following structure. The polymerase fills the gap, adding nucleotides to the exposed 3' end of the nicked chain. *E. coli* pol I has a second enzymatic activity which allows it to digest a polynucleotide chain from




The four nucleotide triphosphate precursors of DNA

the 5' end. Thus, as synthesis proceeds at the 3' side the nick is "translated" as shown in (Eq. 5-24).




(5-24)

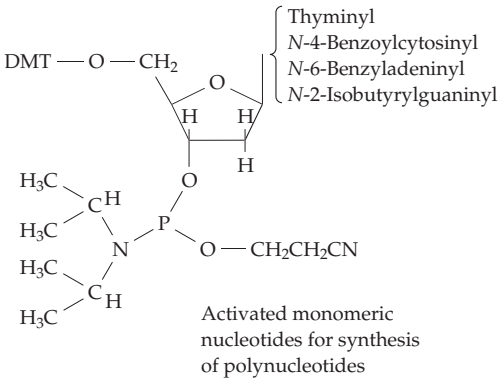
Probes may also consist of DNA copied from mRNA. This is known as **cDNA** and is also widely used to determine indirectly the sequences of mRNA molecules. Messenger RNA may be isolated from the total cellular RNA by affinity chromatography on bound poly (dT) or poly (U). These materials selectively hold RNA with the poly (A) tails characteristic of most eukaryotic mRNA (see Chapter 28). Another source of mRNA is polyribosomes (polysomes), which are "reading" mRNA and actively making proteins. An antibody to the protein for which mRNA is desired will often bind to the protein chains being synthesized and precipitate the polysomes. The mRNA can be recovered and used as a template for cDNA.

Synthesis of cDNA, usually in radiolabeled form is accomplished with **reverse transcriptase**, the enzyme from retroviruses that synthesize a DNA–RNA hybrid from ssRNA.[570–572] A short oligo (dT) primer is usually hybridized to the 3' poly (A) tail to initiate synthesis. Reverse transcriptase also has ribonuclease (RNase H) activity and will digest away the RNA. If desired, synthesis of the second strand can be carried out by a DNA polymerase to give a complete DNA duplex. Many gene sequences have been deduced from cDNA copies.

Often the most practical approach to obtaining a DNA probe is synthesis of a mixture of short oligonucleotides, often in radioactive form as described in the next section. The "redundancy" in the genetic code, i.e., the existence of two or more codons for most amino acids, presents a problem in designing an oligonucleotide probe based on amino acid sequence information. Examination of Table 5-5 suggests part of the solution. Whereas only Met and Trp have single unambiguous codons nine amino acids have only two codons each. We should try to find an amino sequence that contains Met and Trp and as many of the nine others as possible. We should avoid sequences that contain Ser, Leu, or Arg because each has six codons. We can then make a mixture of oligonucleotides, using

the various coding possibilities. Mixtures of as many as 1024 ($2^{10}$) oligonucleotides have been used. One of these may bind tightly and specifically to the desired DNA segment. Instead of such a complex mixture it may be more useful to incorporate a modified base at the most ambiguous positions. For example, inosine, which occurs in the wobble position in anticodons (Fig. 5-30), can pair with A, C, or T.[573] Substitution of 2-aminoadenosine can cause a probe to bind more tightly because a third hydrogen bond will be present in each AT pair.[574]

Another important procedure is labeling ends of polynucleotides. Most often the 5' end is labeled with a radioisotope or by covalent attachment of a fluorescent dye. For example, a **polynucleotide kinase** can be used to transfer a radioactive γ-phospho group from ATP to the 5' end of a polynucleotide that has a free 5'-OH group.



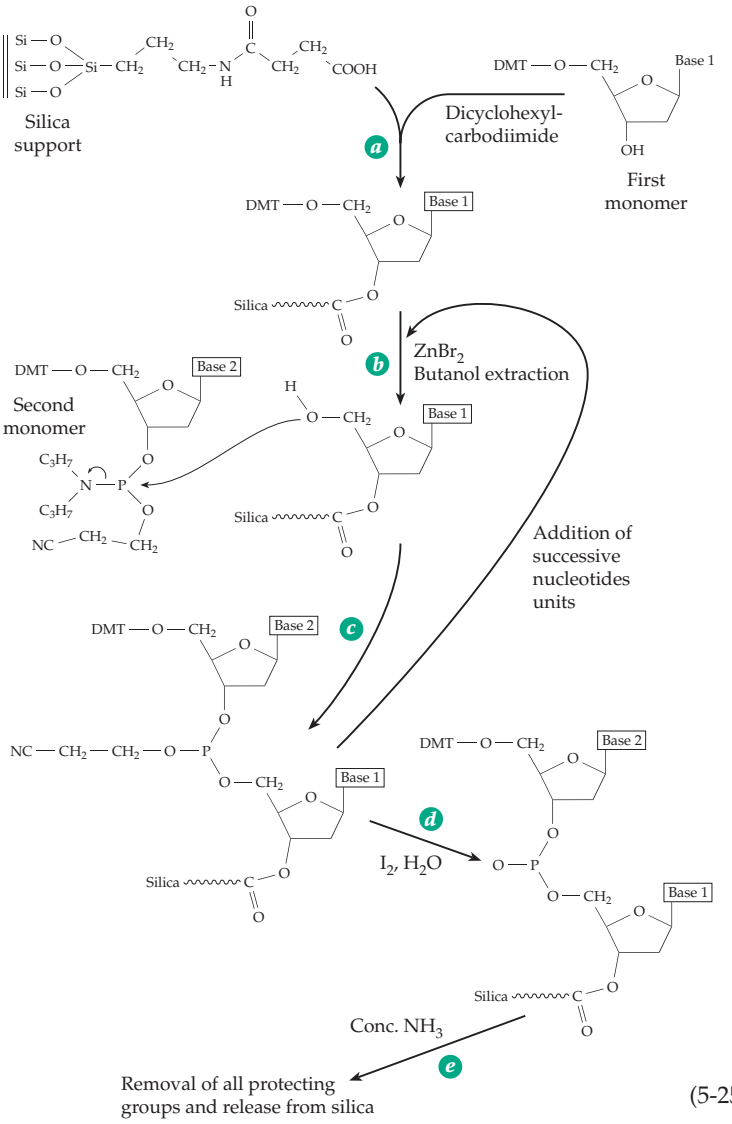
Activated monomeric nucleotides for synthesis of polynucleotides

## 5. Synthesis of Oligonucleotides and Polynucleotides

Efficient solid-phase methods of synthesis analogous to those for polypeptides (Fig. 3-15) have been devised. The pioneering work was done by H. G. Khorana, who made the first synthetic gene[575] and later synthesized a gene for the visual pigment rhodopsin (Chapter 23). Several synthetic approaches have been developed.[576] Currently the most popular method involves the use of phosphite esters. Most nucleophilic groups of the monomers are derivatized with removable blocking groups. For example, N-4 of cytosine and N-6 of adenosine may carry benzoyl groups. The 5'–OH of each nucleotide is blocked by a di-*p*-anisylphenylphenylmethyl (also called dimethoxytrityl, DMTO) group. The 3'–OH is converted to one of a number of activated derivatives such as the following *N*, *N*-diisopropylamino phosphines.[575,577–579]

Solid-phase synthesis is usually done on a silica support with a covalently attached succinamide as shown in Eq. 5-25. The first nucleotide at the 3' end of the chain to be synthesized is attached by an ester linkage to the bound succinamide (step *a*, Eq. 5-25). The 5'-protecting group is removed in step *b* and the 5'-OH reacts with the activated phosphine of the second nucleotide (step *c*, Eq. 5-25). Steps *b* and *c* are then repeated as often as necessary to complete the chain. The finished polynucleotide can be removed from the solid support, the cyanoethyl groups removed from the phosphorus atoms by β elimination and all of the other blocking groups removed by treatment with concentrated $NH_3$. The whole procedure has been automated.[580–582]

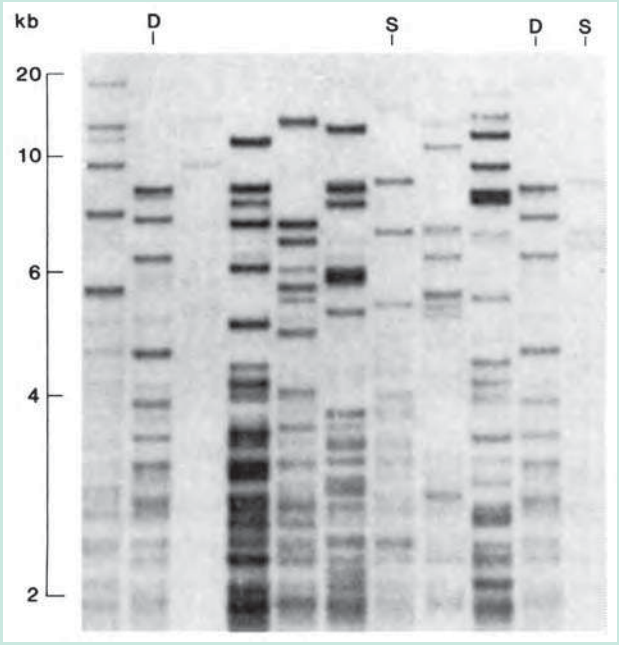
If a large piece of DNA is needed several oligonucleotides can be joined end to end enzymatically (Eq.




(5-25)

## BOX 5-D DNA FINGERPRINTING

Jeffreys *et al.*[a–c] digested human DNA to completion with *Hin* f I and *Sau*3A restriction endonucleases. Certain fragments, which originated from "minisatellite" bands of repetitive DNA showed a very high degree of **polymorphism** among the population. Many different fragments sharing these repeated sequences were formed in the restriction digest. If a suitable labeled probe was used, it hybridized with as many as 80 different bands.[a,d–f] The resulting pattern appeared, like a fingerprint, to be different for every individual, as is shown in the accompanying photo. Unlike a fingerprint the DNA pattern also contains information that often allows deductions about parentage.




DNA "fingerprints" made from one or two drops of blood from ten different individuals. DNA was isolated, digested to completion with restriction endonuclease *Hin* f I, and subjected to electrophoresis in a 20-cm-long agarose gel until all DNA fragments smaller than 1.5 kb in length had passed off the gel. The DNA was then transferred to a nitrocellulose filter by Southern's method and hybridized with a $^{32}$P-labeled single-stranded DNA probe prepared from cloned human minisatellite DNA. The probe used had the "consensus" composition (AGAGGTGGGCAGGTGG). Within the 29 tandem repeats in this 0.46-kb probe there are various sequences close to the one shown. Filters were then autoradiographed for four days. Two duplicate samples (marked D) were taken from the same individual and two others (marked S) from two sisters. A number of bands in common are evident. From Jeffreys *et al.*[c]

The technique has come into widespread use in forensic analysis, with DNA typing being possible from a single hair.[g] In a famous early case it was used to allow an immigrant child to be reunited with his mother[h] and it is being used regularly to protect innocent persons accused of rape or murder.[i] It is also widely used to provide evidence of guilt. However, the very small chance of a close match between unrelated persons prevents the use of DNA typing alone as proof of guilt. Because DNA samples are often "amplified" by PCR (Section H,6), there is also a possibility of contamination and forensic use of DNA typing is still controversial.[i–m]

However, DNA typing continues to be improved[f,n,o] and to be applied in a great variety of ways. For example, the skeletal remains of a murder victim were identified by DNA fingerprints after being buried for eight years.[p] DNA typing is also used to study mating habits of birds,[q] the genetic variability of populations of whales sampled by biopsy,[r] etc.


[a] Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985) *Nature (London)* **314**, 67–73
[b] Lewin, R. (1986) *Science* **233**, 521–522
[c] Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985) *Nature (London)* **316**, 76–79
[d] Vassart, G., Georges, M., Monsieur, R., Brocas, H., Lequarre, A. S., and Christophe, D. (1987) *Science* **235**, 683–684
[e] Huang, L.-S., and Breslow, J. L. (1987) *J. Biol. Chem.* **262**, 8952–8955
[f] Kirby, L. T. (1990) *DNA Fingerprinting*, Stockton Press, New York
[g] Higuchi, R., von Beroldingen, C. H., Sensabaugh, G. F., and Erlich, H. A. (1988) *Nature (London)* **332**, 543–546
[h] Jeffreys, A. J., Brookfield, J. F. Y., and Semeonoff, R. (1985) *Nature (London)* **317**, 818 – 819
[i] Balding, D. J., and Donnelly, P. (1994) *Nature (London)* **368**, 285–286
[j] Neufeld, P. J., and Colman, N. (1990) *Sci. Am.* **262**(May), 46–53
[k] Lewontin, R. C., and Hartl, D. L. (1991) *Science* **254**, 1745–1750
[l] Lander, E. S., and Budowle, B. (1994) *Nature (London)* **371**, 735–738
[m] Lewontin, R. C. (1994) *Nature (London)* **372**, 398
[n] Uitterlinden, A. G., Slagboom, P. E., Knook, D. L., and Vijg, J. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2742–2746
[o] Jeffreys, A. J., MacLeod, A., Tamaki, K., Neil, D. L., and Monckton, D. G. (1991) *Nature (London)* **354**, 204–209
[p] Hagelberg, E., Gray, I. C., and Jeffreys, A. J. (1991) *Nature (London)* **352**, 427 – 429
[q] Burke, T., and Bruford, M. W. (1987) *Nature (London)* **327**, 149–152
[r] Hoelzel, A. R., and Amos, W. (1988) *Nature (London)* **333**, 305

27-5). For example, a functional 17-bp gene for the 53-residue human epidermal growth factor was synthesized by joining ten oligonucleotides of lengths 11–59 bp.[575] DNA is often synthesized enzymatically using methods described in Chapter 26. Cloned sequences of synthetic DNA can also be transcribed to produce **polyribonucleotides** of any desired sequence.[583] New nonenzymatic methods for RNA synthesis have also been devised.[583–587]

## 6. The Polymerase Chain Reaction (PCR)

This important technique was first described in 1971–1974 by Khorana and associates[588,589] but was not used until it was rediscovered in 1983 by Mullis.[590–593] It was quickly developed[591,592,594–596] and has played a major role in biochemistry ever since. It continues to be applied in numerous ways.[589,597–599a]

The PCR technique provides a way of "amplifying" a small number of DNA molecules, i.e., to produce many copies. This is often done by cloning but PCR offers a quick and easy way to obtain millions of copies of a desired relatively short segment of DNA. Standard PCR can be used for up to about 5000-nucleotide pieces. More recently modified procedures have allowed 35-kb segments to be amplified.[600]

The basic PCR procedure is initiated by hybridizing two oligonucleotide primers onto opposite strands of denatured DNA, one at each end of the section chosen for amplification (Fig. 5-47). A DNA polymerase is then used to convert each of the separated strands into a duplex. The mixture of products is then heated to denature the two new duplexes. After cooling, the primers, which are present in great excess, hybridize to all four strands. In a second cycle of polymerase action these are all converted to duplexes, etc. After 20 cycles millions of copies will be made. At first, copies with tails extending beyond the limits specified by the oligonucleotide primers will be formed. However, it is easy to see that after a few cycles, most molecules will be of just the desired length. A heat-stable polymerase from *Thermus aquaticus* (*Taq* polymerase) is used so that the enzyme is not denatured by the repeated cycles of heating and cooling, which are conducted automatically by a simple apparatus.

The polymerase chain reaction is being used to speed up prenatal diagnosis of genetic diseases, to detect viral infections, for tissue typing needed for organ transplantation, in forensic procedures, and in the study of the DNA of ancient tissues such as those of frozen wooly mammoths.[601–603] If suitable restriction enzyme sites are present in the primers, the amplified DNA can be cloned readily.[604] The $3.3 \times 10^{-9}$ fmol of a DNA sequence found in a diploid chromosome pair in a single cell can be amplified in 50 cycles to 5–500 fmol, enough to study by hybridization with radioactive probes.[605] A large sample of a few pg of DNA can be amplified in 20 cycles to micrograms. By placing sequencing primers within the amplified segments, it is possible to generate DNA that can be sequenced directly using the dideoxy sequencing technique (Chapter 5) without cloning.[606] The PCR technique has also been used to amplify cDNA molecules formed from RNA transcripts present in very low abundance. One of the problems with the PCR is that priming may occur by DNA fragments other than the added primers. Contamination must be scrupulously avoided. Another problem is that errors are introduced into DNA during amplification by PCR. Perhaps 1 in 200 of the copies will contain an incorrect base.[607] If such a molecule is cloned the error will be perpetuated. Good practice requires that more than one clone is selected and sequenced to allow such errors to be avoided.

## 7. Sequence Determination

Satisfactory (but slow) methods for determining sequences of RNA molecules have been known for over 30 years. The procedures are somewhat parallel to those used in sequencing proteins. However, no similar method could be devised for DNA. Little progress was made until rather recently when new approaches led to extremely rapid procedures for sequencing DNA. As a consequence, it is now much easier to learn the sequences of genes than it is to sequence the proteins which they encode!

***Preparing the DNA.*** The first step is to obtain a sample of enough identical DNA molecules to permit sequence analysis. This in itself may be a complex undertaking. Perhaps we want to know the sequence of one particular gene in the 3,500,000 kilobase pairs of DNA present in a single human cell. How can this gene be found and the DNA be obtained for analysis? Three techniques have been essential: cutting the DNA with restriction endonucleases, hybridization, and cloning. More recently PCR and related methods[608] have simplified the sample preparation. DNA can often be amplified using primers that contain sequences that will later serve as **sequencing primers**.

***Restriction maps and Southern blots.*** Although it doesn't require the synthesis of a primer, the Maxam–Gilbert procedure usually demands that a "restriction map" of the DNA be prepared to help keep track of the fragments being sequenced.[609] See also Chapter 26. Figure 5-48[610–612] shows the restriction map of the mitochondrial DNA gene, *oxi3* from yeast. This gene, which encodes one of the subunits of cytochrome oxidase (Chapter 18), consists of 9979 base pairs. It was cloned in a suitable plasmid after which the restriction map (Fig. 5-48) was prepared by cutting with 18 different

restriction enzymes.[611] The protein subunit contains 510 residues and therefore requires a coding capacity in the DNA of 1530 base pairs. This is only 16% of the total length of the gene, the majority of whose DNA is found in the three large introns.

One way to select a desired segment of DNA from a digest of chromosomal DNA is to sort out the "restriction fragments" by gel electrophoresis.[613] The DNA from the gel can be transferred to a nitrocellulose sheet while retaining the separation pattern using a method devised by Southern.[560,614,615] In this **Southern blot** technique, solvent flows from a pool beneath the gel up through the gel and the nitrocellulose sheet into paper towels. The DNA is trapped on the nitrocellulose in the same pattern observed in the electropherogram. A suitably labeled probe such as cDNA with incorporated $^{32}P$ is flowed repeatedly across the nitrocellulose sheet under conditions that favor formation of hybrids. Only the DNA complementary to the cDNA probe will retain the label. This DNA can then be located with the help of a autoradiogram. It is important that single-stranded DNA be used. If double-stranded restriction fragments are separated on the electropherogram they must be denatured while in place in the gel before hybridization is attempted.

Once the desired piece of DNA has been identified it is usually necessary to increase its amount. The conventional approach is to incorporate the DNA fragment into a plasmid and clone by the methods described in Chapter 26. The selected DNA can usually be cut cleanly from the plasmid used for cloning with the same restriction endonuclease originally used in

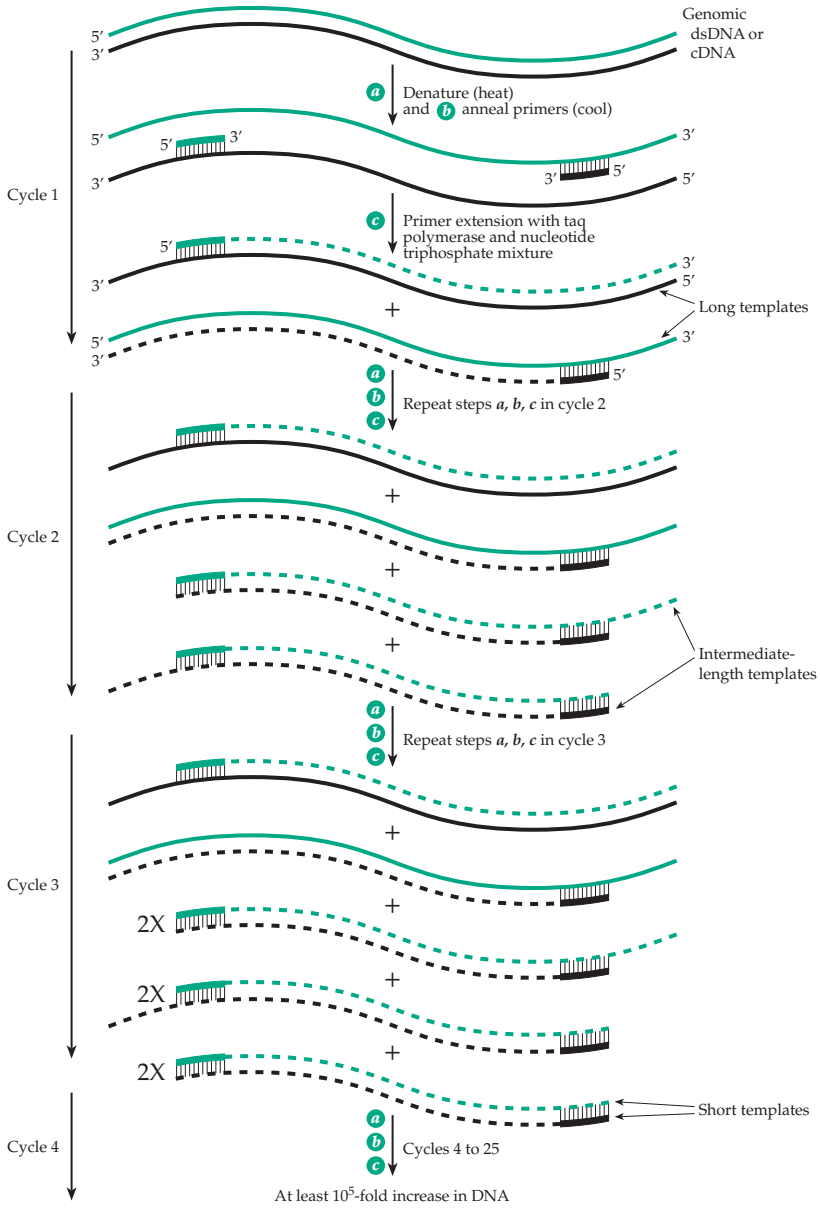


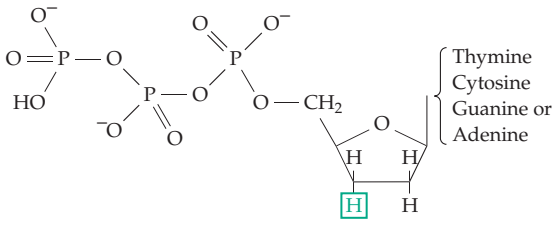

**Figure 5-47** Amplification of DNA using the polymerase chain reaction (PCR). Double-stranded DNA is denatured by heating to 90–99° C (step *a*) and oligonucleotide primers complementary to short 12–18 nucleotide sequences at the two ends of the piece of DNA to be amplified are annealed to the separated strands by cooling to 40–75° C (step *b*). The two DNA strands serve as templates for synthesis of new complementary strands using a heat-stable DNA polymerase and a mixture of the four nucleotide triphosphates. Nucleotide units are added to the 3' ends of the primers, with the new chains growing in the 5' → 3' direction (step *c*). Steps *a*, *b*, and *c* are then repeated as many as 30 times using a thermal cycler device that periodically raises and lowers the temperature with a cycle time of a few minutes. The polymerase is unharmed by the heating and is reused in each cycle. An excess of the primer and of the nucleotide triphosphates sufficient for all of the cycles is present initially. In the early cycles new long and intermediate length templates are created. However, the number of short templates increases exponentially and the final product consists predominately of the short selected DNA segment (short templates).

fragmenting the DNA. An alternative procedure is to clone a mixture of DNA fragments and then sort colonies of bacteria containing the cloned fragments using DNA–RNA hybridization.[616] Bacteria from the selected colonies are then propagated to produce large amounts of the plasmid DNA. Alternatively, PCR can be used directly on the selected DNA fragment. This is more often the preferred choice.[604,606]

***The Sanger dideoxy method.*** The rapid sequencing methods all depend upon the fact that single-stranded DNA fragments under denaturing conditions migrate on electrophoresis in polyacrylamide gels strictly according to their length. Thus, if a mixture contains all lengths of radiolabeled polynucleotides from very short oligonucleotides to fragments containing 200 or 300 bases, the polynucleotides will all appear, one above the other, as a series of bands that can be visualized by radioautography. The first of these methods was published by Sanger and Coulson[617] in 1975 and was followed in 1977 by the method which is now used.[618–621]

A sample of double-stranded DNA is denatured. One of the resulting single strands is used as a template to direct the synthesis of a complementary strand of radioactive DNA using a suitable DNA polymerase. The "Klenow fragment" of *E. coli*, DNA polymerase I, reverse transcriptase from a retrovirus, bacteriophage T7 DNA polymerase, *Taq* polymerase, and specially engineered enzymes produced from cloned genes have all been used.

Before the sequencing begins it is necessary to prepare a short **primer** that is complementary to a sequence at one end of the DNA strand to be sequenced. This may be prepared enzymatically,[622,623] or by nonenzymatic synthesis. The short primer is annealed to the end of the DNA and the resulting molecule is incubated with a DNA polymerase and a mixture of the four mononucleotide triphosphates, one of which is radiolabeled in this position. Four reaction mixtures are prepared. Each mixture contains all four nucleoside triphosphates and also one of four different **chain-terminating inhibitors**, the most popular of which are the 2', 3'–dideoxyribonucleoside triphosphates:




2',3'-Dideoxyribonucleotide triphosphates

These inhibitors are added in a ratio of about 100:1 with the natural substrates. In this ratio they are incorporated into the growing DNA chain about once in 200 times on the average. However, in the various growing DNA chains they are incorporated at different points ranging from the very first nucleotide to the last. Since the incorporated dideoxy monomer lacks the 3-hydroxyl group needed for polymer formation, chain growth is terminated abruptly. Synthesized polynucleotides are denatured and subjected to electrophoresis in four adjacent lanes. The resulting patterns contain bands corresponding to all of the successive oligonucleotides but not all in the same lane. A given lane will contain only the bands of the oligonucleotides terminated by the particular inhibitor used. The other bands will be found in the other three lanes. Each band will have been terminated by the inhibitor employed in that lane. The nucleotide sequence can be read directly from the banding pattern as is shown in Fig. 5-49. Arabinosyl nucleotide triphosphates have also been used as chain-terminating inhibitors. A sequence determined by the Sanger method is usually checked by also sequencing the complementary strand.

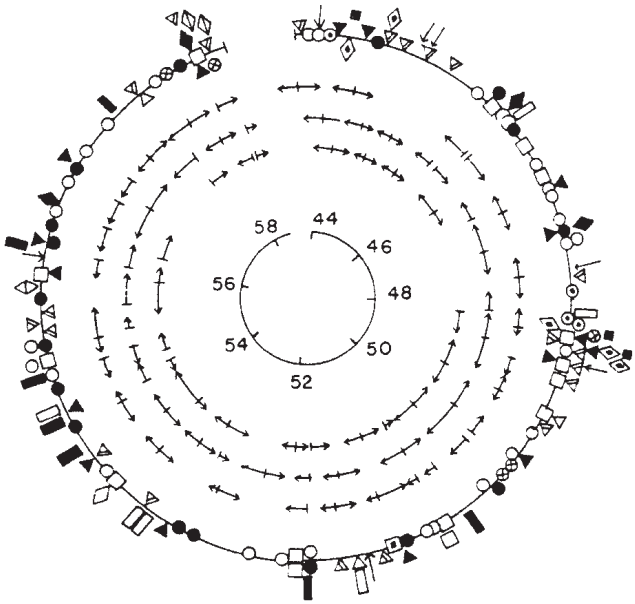



**Figure 5-48** A physical map of the *oxi3* locus of yeast mitochondrial DNA. The restriction fragments used for DNA sequencing are indicated by the *arrows*. The extent to which the sequences were read is represented by the lengths of the arrows. The map units are shown in the *inner circle*. The following symbols, together with the names of the restriction enzymes (Chapter 26), are used for the restriction sites:

| | | |
|---|---|---|
| ▲ *Hinf* I | ○ *Alu* I | ⊙ *Hha* I |
| △ *Hpa* II | ◈ *Pvu* II | ◬ *Rsa* I |
| ▯ *Hae* III | ◈ *Hinc* II | ▮ *Hph* I |
| ◆ *Taq* I | ◈ *Hind* III | ⊡ *Blg* II |
| ● *Mbo* I | ⊗ *Eco* RI | ◇ *Bam* HI |
| □ *Mbo* II | ↓ *Eco* RII | ■ *Hpa* I |

From Bonitz *et al.*[611]

A major factor in the success of the dideoxy sequencing method has been the development of cloning techniques that provide ssDNA in a form ready for use. In any cloning procedure the DNA that is to be sequenced has been covalently ligated to the end of a DNA strand of the **cloning vehicle**, a modified plasmid or virus (Chapter 26). Sequencing is often done on DNA cloned in a modified ssDNA bacterial virus such as M13. Although dsDNA is ligated to the ds replicating form of viral DNA, the virus particles produced when the virus is propagated in *E. coli* cells are single stranded. Cleavage of the viral DNA with its incorporated "passenger DNA" with an appropriate restriction endonuclease releases the passenger DNA (to be sequenced) with a short piece of DNA from the cloning vehicle attached at the 3' end. Since the sequence of this small piece of the cloning vehicle is known, a suitable primer of length ~12–18 nucleotides can be synthesized (or purchased) and annealed to the DNA. This serves as the primer and allows the sequence to be read from the 5' end to the 3' end of the synthesized complementary strand. Double-stranded DNA attached to vehicles such as the pUC plasmids can also be sequenced directly if the DNA is denatured by alkali treatment. After neutralization and precipitation an appropriate primer is annealed to one or the other of the two strands.[624]

Since its introduction, many modifications and improvements have been developed. The sequencing gels have been improved. The use of $^{35}S$ labeling has given sharper autoradiographs.[625] Alternatively, a silver stain can be used with unlabeled primers.[626] GC-rich DNA sequences are often difficult to sequence, probably because even in the denaturing polyacrylamide gels used for sequencing they tend to form hairpin loops, perhaps as a result of formation of Hoogsteen base pairs (Fig. 5-7). Formation of these loops results in uneven spacing between the adjacent bands in the sequencing gel, so-called "compression artifacts." Use of a 7-deaza-dGTP in place of dGTP in the sequencing reaction ameliorates this problem.[627] Sensitivity can be improved by use of "cycle sequencing" in which a heat-stable polymerase such as *Taq* polymerase is used, and after heating the same template DNA is used repeatedly to give a higher yield of labeled fragments.[628,629]

About 200 – 400 bases can be successfully sequenced manually in a single run. By prolonging the time of electrophoresis in a second run, the sequence can be extended considerably. By using the just obtained sequence information, it is possible to select a new start point 200 or more nucleotides further along the template chain and to synthesize an oligonucleotide primer to anneal to the template at this point. In this way it is possible to "walk" along the template adding additional sequences at each step. Another procedure is to delete by mutation various segments of the cloned DNA above the sequence that binds the primer. This

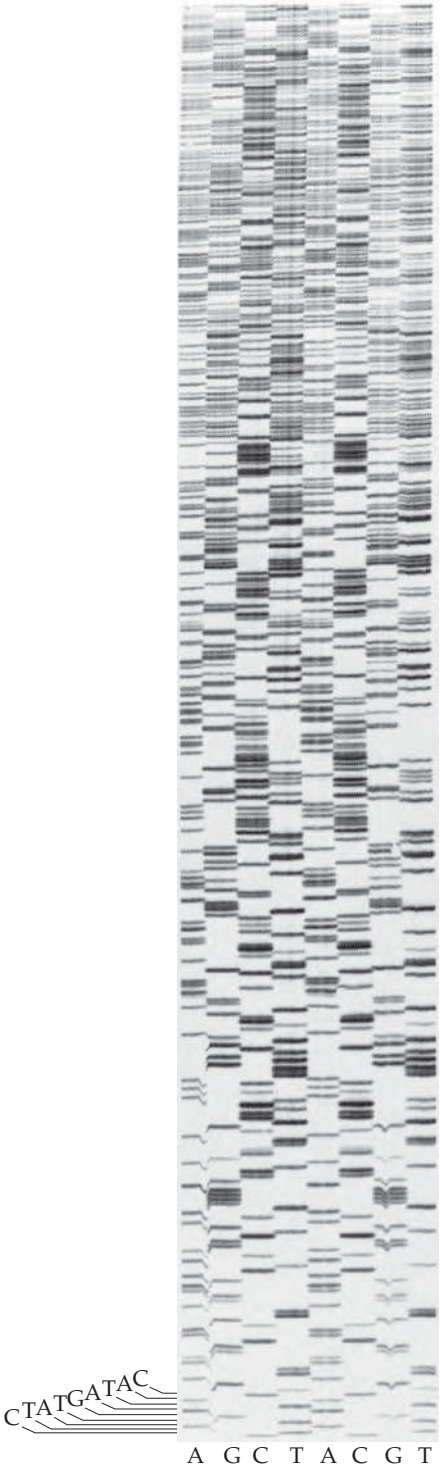



**Figure 5-49** A DNA sequencing gel obtained using a segment of DNA from salmon sperm selected by suitable oligonucleotide primers, amplified by PCR, and sequenced with a $^{35}S$ label in the primer. Four samples were used, one with each of the four dideoxy chain terminators (A, G, C, T, A, C, G, T from left to right). After electrophoresis the shorter fragments are at the lower end of the gel. The sequence of the strand complementary to the template strand whose sequence is being determined is read from the bottom of the gel. Here it starts CTATGATAC. Reproduced by permission of Amersham Pharmacia Biotech, Limited.

permits analysis of the whole cloned fragment via an overlapping set of sequences from the deletion mutants.[630,631]

Many DNA sequences continue to be determined manually by the well-developed long gel procedures as illustrated in Fig. 5-49. However, sequencing whole genomes has depended upon the development of high-speed automated procedures.[632] Instead of radiolabeling, fluorescent dyes may be joined to the primer to allow detection of the chain fragments produced during sequencing. Automatic sequencers use four dyes that fluoresce with different colors.[633,634] A different dye is used for each of the reaction mixtures. Then the four samples are mixed together and the DNA fragments are subjected to electrophoresis in a single lane. A laser beam excites the fluorescence, scanning several lanes with different samples as the electrophoresis progresses. A photomultiplier tube records the fluorescence intensity of each band through a series of four filters in a rotating wheel. This allows automatic recognition of the four different colors of fluorescence and therefore of the nucleic acid base present in each position in the sequence. Improved strategies for "primer walking"[635] and for "shotgun sequencing"[19,632,636] have been devised. In the shotgun strategy, whole bacterial genomes have been cut by restriction enzymes into large numbers of overlapping fragments which have been separated and sequenced.[637] A computer program is used to analyze and assemble the sequences into a complete genomic sequence. An example is provided by the genome of the *Methanococcus jannaschii*. Its large circular chromosome contains 1,664,976 bp and there are two additional pieces to the genome, one containing 58,407 bp and the other 16,550 bp. The sequences were deduced from 36,718 individual sequencing runs on high-speed automatic sequencers. For each run, on average, 481 bp could be read.[636] To sequence the human genome faster methods are needed.[638] Capillary electrophoresis with a single laser beam scanning the output of 24 capillaries has been demonstrated.[639,640] Extremely rapid sequencing of oligonucleotides up to 100 bp in length can be accomplished by mass spectrometry.[641] This may be an important technique for diagnosis of genetic defects (Chapter 26).

***The method of Maxam and Gilbert.*** The nonenzymatic method devised by Maxam and Gilbert[642–644] can be used to sequence either ss or dsDNA. Before the sequencing is begun, a radioactive label is incorporated, usually at the 5' end. This is often done by cleaving off any phosphate groups present on the 5' end with alkaline phosphatase and then transferring a new radioactive phospho group with the assistance of the enzyme polynucleotide kinase and radioactive $\gamma$-$^{32}$p-labeled ATP. If dsDNA is used the strands are separated so that each has a label only at one end.

The key step in sequencing by the Maxam–Gilbert procedure is to cleave chemically the DNA at random locations using reagents that have some specificity for particular bases. The cleavage process involves three distinct steps: (1) chemical modification, as specific as is possible for the chosen base; (2) displacement of the modified base from the sugar; and (3) elimination and chain cleavage using amine catalysis (Eq. 5-21). Two consecutive steps can often be combined. There are several versions of the method; one involves dimethyl sulfate as the specific reagent for guanine. It forms $N^7$-methylguanosine (see Eq. 5-18) which upon heating with the strong base piperidine at 90°C undergoes addition of hydroxyl ion with ring opening and displacement of the modified base according to Eq. 5-26. The product, a glycosylamine of piperidine, is in equilibrium with a Schiff base which can undergo chain cleavage as in Eq. 5-21.

A second sample of DNA is treated with a piperidine-formate buffer of pH 2 in the cold. This promotes the acid depurination of both guanine and adenosine. A third sample is treated with hydrazine, with both cytidine and thymidine being cleaved to ribosylurea according to Eq. 5-13. Again, this is followed by displacement and $\beta$ elimination (Eq. 5-21) catalyzed by piperidine. The fourth sample is also treated with hydrazine but in the presence of a high concentration of NaCl which inhibits the reaction with thymidine, by lowering the p$K_a$ of the thymine, and allows the cleavage to be more nearly specific for cytidine. Each of the reactions is conducted in such a way that on the average only one cleavage event occurs per molecule of DNA. Since the cleavages occur at many different points, a family of nested radioactively labeled oligonucleotides, one from each original molecule, is produced. When these are sorted by polyacrylamide gel electrophoresis, the pattern of the oligonucleotides in the four adjacent channels allows the nucleotide sequence to be read directly from the autoradiogram.

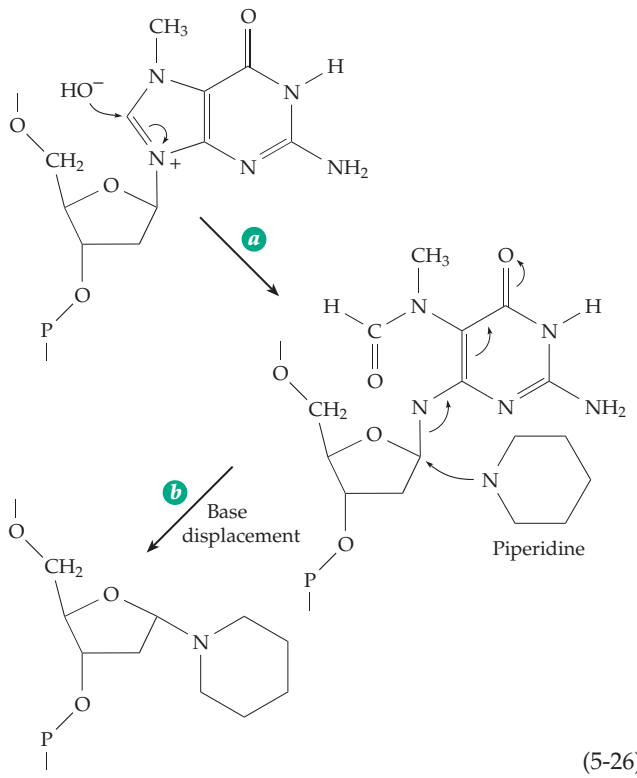
The Maxam–Gilbert method doesn't require synthesis of a primer and it sometimes works well for sequences that are difficult to obtain with the Sanger–Coulson procedure. The two methods may both be used to provide additional certainty about a sequence. The Maxam–Gilbert method is very convenient for sequencing small oligonucleotides which often react poorly with the polymerase used for the chain termination method. The method usually requires that a restriction map be prepared.

***Sequencing RNA.*** The first known RNA sequence, that of an alanine tRNA, was determined by Holley and associates in 1965. The RNA was subjected to partial hydrolysis with pancreatic ribonuclease and ribonuclease $T_1$ (Fig. 5-43). The small oligonucleotide fragments were separated by ion exchange chromatography under denaturing conditions (7 M urea) and were then characterized individually.[645] The availability

of additional enzymes such as ribonuclease $U_2$, the *B cereus* ribonuclease and ribonuclease Phy M of *Physarum* (Fig. 5-43) and the use of radiolabeling and of two-dimensional fingerprinting of digests have made the procedures more versatile.[646] A valine tRNA was sequenced independently by Bayev[647] and Campbell.[648]

Since the development of the rapid methods for sequencing DNA, many mRNA sequences have been determined by using reverse transcriptase to make a cDNA strand complementary to the RNA. The cDNA is then sequenced.[649] Rapid sequencing methods parallel to those used for DNA have also been devised.[650–652]

***Nearest neighbor analysis.*** A technique developed by Kornberg and associates before the availability of sequencing methods is the **nearest neighbor sequence analysis**. Using a single radioactive $^{32}P$-containing nucleoside triphosphate together with the three other unlabeled nucleoside triphosphates, a primer chain of DNA is elongated from the 3' end along a ssDNA template chain using a DNA polymerase. The incorporation of $^{32}P$ from the α position of the nucleotide triphosphate occurs in the bridge phosphates that connect the nucleotide originally carrying the $^{32}P$ to the 3' position of the neighboring nucleotide. Cleavage of the $^{32}P$-containing product of the reaction with a mixture of micrococcal DNase and spleen phosphodiesterase, which catalyze *b*-type cleavage (Fig. 5-43), gives fragments in which the $^{32}P$ will now be attached to what was the 5′ nearest neighbor to the radioactive




(5-26)

nucleotide in the DNA.[72,653] Measurement of the radioactivity in each of the 3' nucleotides of thymine, cytosine, adenine, and guanine gives the frequencies of the adjacent pairs, TA, CA, AA, and GA. Using the other radioactive nucleoside triphosphates one at a time in separate experiments, all of the nearest neighbor frequencies can be obtained. From such an experiment it was possible to deduce that the strands in the double helix were oriented in an antiparallel fashion, as predicted by Watson and Crick. If the strands had been parallel, different nearest neighbor frequencies would have been observed.

***Understanding sequences.*** Sequences of over 20 million nucleotides from hundreds of organisms had been determined by 1988 and the number is doubling each 2–3 years.[654] Sequencing the human genome has required rates of millions of bases per day. With the massive amount of data already available it has become of great interest to compare sequences of genes, whether they encode similar or dissimilar proteins, to make comparisons between species, and to search for sequences that bind specific proteins or that encode particular regulatory signals. Relationships of common evolutionary origin or homology as well as other sequence similarities are often sought.

To handle the mass of existing data, powerful computer programs have been developed and various graphical procedures have also been developed to help the human mind comprehend the results.[654,655] One important problem is to define and locate what are called **consensus sequences**. The problem is best illustrated by examples.[654] The cleavage site for the *Eco*RI restriction endonuclease is **GAATTC**. There is no ambiguity. In a DNA of random sequence this would be expected to occur by chance in about $(1/4)^6$ nucleotides (4 kb). On the other hand, the *Hin*II restriction endonuclease cleaves within the consensus sequence **GTYRAC** where Y = C or T and R = A or G. It would be expected to occur by chance in about $1/4^5$ nucleotides. Many binding sites for RNA polymerase, the so called **promoters** (Chapter 28) contain the consensus sequence **TAtAaT**, at position -10, ahead of the 5' end of the sequence that is transcribed into mRNA. The lower case t and a used here imply that other nucleotides may often replace T or A at these positions. There are many promoters and over 70% of those described have this consensus sequence. All have the less restricted sequence **TAxxxT**, where x may be any nucleotide. Our definition of consensus sequence is somewhat arbitrary. Now consider the problem of locating a –35 site whose consensus sequence is **TTGACA** but which may, for different genes, be shifted backward or forward by a nucleotide or two. This is a consensus sequence. Therefore, in many cases one or more substitutions in the sequence will have been made. The result is that the sequence of nucleotides in

which the consensus sequence is to be found is likely to appear entirely random. Sophisticated computer programs are helpful in locating it.[654]

## 8. Protein–DNA Interactions

The most detailed information about interactions of proteins with DNA is coming from X-ray crystallographic studies. Examples are seen in Figs. 5-35 to 5-40. Several other methods have also been very useful. Much has been learned from the effects of mutations in DNA-binding proteins or in regions of DNA to which a protein binds. Binding of proteins to DNA can also be recognized by its effects on the mobility of DNA during gel electrophoresis.[656,657] Chemical[658] or laser-induced crosslinking can show that within a complex a specific residue in a protein is adjacent to a certain sequence in the DNA.

The technique of **protection mapping** or **"footprinting"** is widely used to determine which nucleotides in a sequence are covered by a bound protein.[659] A reagent which attacks and cleaves DNA nearly randomly is used. DNase I was first introduced for this purpose[659] and has been used widely. An important finding is that certain sites that are readily cleaved (**hypersensitive sites**) are frequently located in chromatin undergoing transcription. Footprinting has also been accomplished with other nucleases, with dimethylsulfate (which acts on A and C), with carbodiimides (which act on U and G), and with the Maxam–Gilbert guanine-specific cleavage (Eq. 5-26). One of the most popular methods employs cleavage by hydroxyl radicals.[660–663] **Photofootprinting** depends upon decreased or increased sensitivity to ultraviolet light at sites bound by proteins.[214,662,664] In footprinting experiments the DNA to be studied is radioactively labeled at one end of one strand. In the absence of the protecting protein, denaturation and electrophoresis of the cleaved fragments yields a nearly random "ladder" of DNA fragments. In the presence of the binding protein some cleavage products will be missing from the ladder. The bound protein leaves a "footprint" (Fig. 5-50A,C).

Related methods are being applied to the determination of the secondary structure of RNA molecules[665,666] and to the study of interactions with proteins. For example, treatment with dimethyl sulfate under appropriate conditions methylates bases that are not paired, giving largely 1-methyladenosine and 3-methylcytidine.[667]

## 9. Nuclear Magnetic Resonance

Much of the initial effort to study polynucleotides by NMR spectroscopy was directed toward transfer RNAs, only a few of which have been crystallized in a form suitable for X-ray diffraction. Study of the other tRNAs by NMR techniques has established that all of the tRNAs have a similar architecture and that the structures observed in the crystals are preserved in solution.[668] Figure 5-51 shows the low-field end of the NMR spectrum of a valine-specific tRNA from *E. coli.* The spectrum is run in $H_2O$ rather than $D_2O$ so that exchangeable hydrogens in the hydrogen bonds of the Watson–Crick base pairs can be observed.[669] The protons giving rise to the downfield resonances are primarily those attached to nitrogen atoms of the rings and in hydrogen-bonded positions. These protons are shielded by adjacent electron-donating groups and by their attachment to the semiaromatic rings of the bases. The NMR signals are further shifted downfield to varying degrees depending upon whether or not the proton being observed is attached to a base that is stacked with other bases. The size of the shift also depends upon which neighboring bases are present. The proton on N-3 of AU base pairs is deshielded more than the proton on the N-1 of GC base pairs. Therefore, the AU protons appear further downfield than the GC protons. The stronger ring current in A than in C enhances this separation.

All of the resonances in Fig. 5-51 have been assigned to particular bases. This was done in part by varying the temperature, changing the magnesium ion concentration, and predicting shifts caused by ring currents in adjacent bases making use of the X-ray crystal structures. NMR spectra of hairpin helical fragments also provided essential information. From integration of the areas under the peaks it was concluded that the 20 resonances seen below –11 ppm represent 27 protons. Twenty of these are in Watson-Crick base pairs and correspond to those expected from the X-ray structure. Six more belong to protons involved in tertiary interactions, such as base pair triplets or non-Watson–Crick pairs. One of these, labeled "G" in the figure, is in the dihydrouridine stem and involves the ring proton of N-1 of $m^7$G46 which is hydrogen bonded to N-7 of G22 in the major groove of the RNA. G22 is located at the beginning of the "extra loop."

Measurements of the NOE of nearby protons in both small RNA molecules[668,669] and DNA oligonucleotides[670,671] provided much additional information. Figures 5-51B and C show NOESY spectra of the $tRNA^{Val}$ and the way in which weak cross-peaks between the H-bonded imino protons in adjacent base pairs (Fig. 5-51B) can be used to establish connectivities.[669] Beginning with resonance B, it is possible to establish the sequence of the NH groups giving rise to these resonances as OBUGJNT. Using other data as well, it was concluded that these represent the seven base pairs of the acceptor stem (see Fig. 5-30), with resonance C representing the first GC pair, resonance B the second, etc. Resonance A was identified as coming from the base triplet containing a Hoogsteen

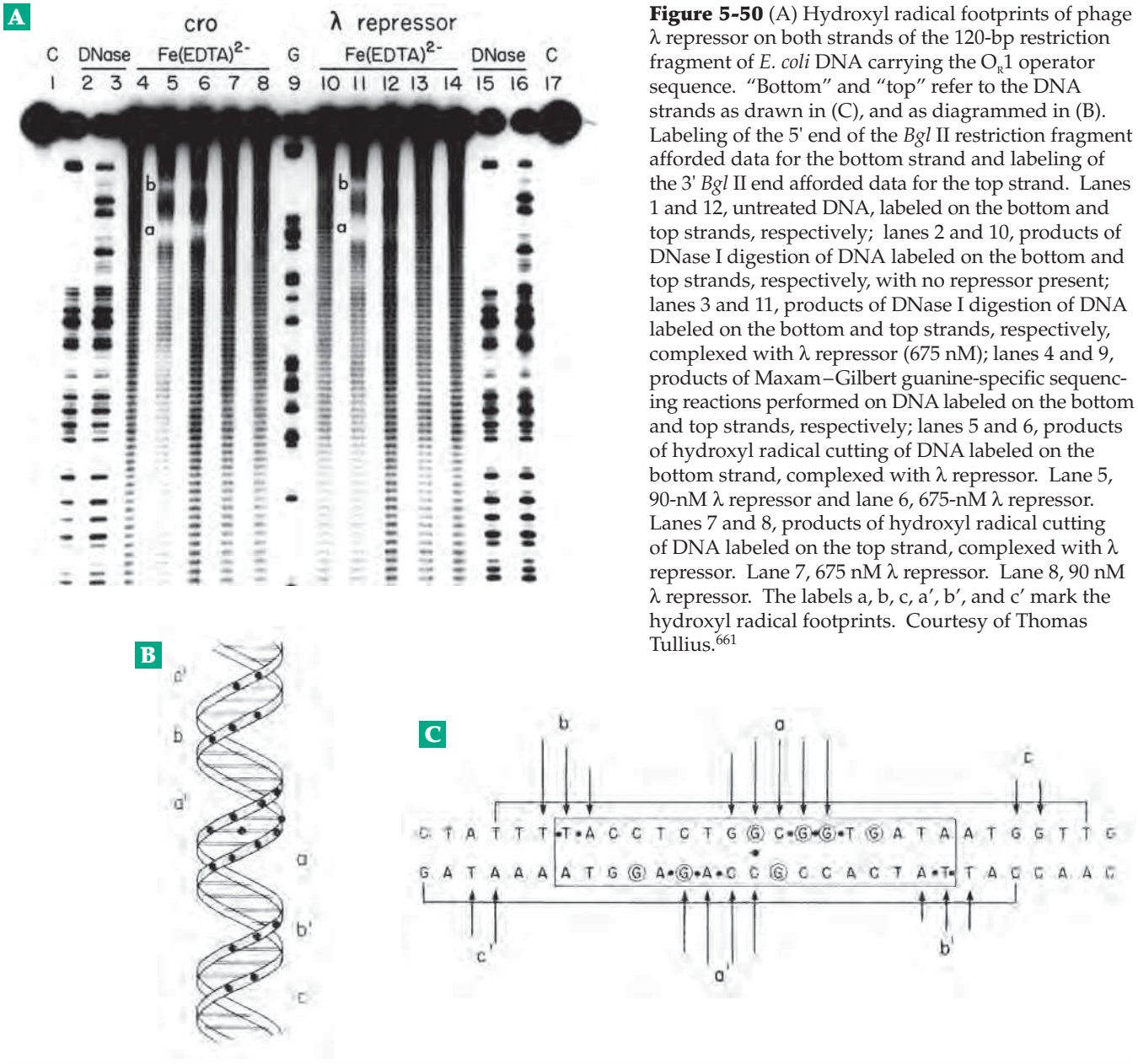



**Figure 5-50** (A) Hydroxyl radical footprints of phage λ repressor on both strands of the 120-bp restriction fragment of *E. coli* DNA carrying the $O_R1$ operator sequence. "Bottom" and "top" refer to the DNA strands as drawn in (C), and as diagrammed in (B). Labeling of the 5' end of the *Bgl* II restriction fragment afforded data for the bottom strand and labeling of the 3' *Bgl* II end afforded data for the top strand. Lanes 1 and 12, untreated DNA, labeled on the bottom and top strands, respectively; lanes 2 and 10, products of DNase I digestion of DNA labeled on the bottom and top strands, respectively, with no repressor present; lanes 3 and 11, products of DNase I digestion of DNA labeled on the bottom and top strands, respectively, complexed with λ repressor (675 nM); lanes 4 and 9, products of Maxam–Gilbert guanine-specific sequencing reactions performed on DNA labeled on the bottom and top strands, respectively; lanes 5 and 6, products of hydroxyl radical cutting of DNA labeled on the bottom strand, complexed with λ repressor. Lane 5, 90-nM λ repressor and lane 6, 675-nM λ repressor. Lanes 7 and 8, products of hydroxyl radical cutting of DNA labeled on the top strand, complexed with λ repressor. Lane 7, 675 nM λ repressor. Lane 8, 90 nM λ repressor. The labels a, b, c, a′, b′, and c′ mark the hydroxyl radical footprints. Courtesy of Thomas Tullius.[661]

base pair of 4-thiouracil at position 8 with A14 (see Fig. 5-7). Its connectivity to the sequence KCEO is also outlined in Fig. 5-51B. However, it could not be established without additional data which also helped to identify the sequence O-K as residues 10–13 of the dihydrouracil stem. Peak O is a multiproton peak representing not only GC 10 but also UA 7. In general, the GC protons are at the higher field side of the spectrum and the AU protons at the lower side. However, AU 7 is shifted to an anomolously high position. Cross-peaks between imino protons of uracil and the nearby C2 protons of adenine in Watson–Crick AU base pairs or C8 protons of Hoogsteen AU pairs can be observed in the 6.5 – 9 ppm region as shown in Fig. 5-35B. This region also contains information about other protons bound to the nuclei acid bases.

Similar techniques are being used for the study of DNA.[672] The presence of a second hydrogen in the 2' position of the deoxyribose rings of DNA adds several H-H distances (Fig. 5-52) that can be measured in addition to those seen in RNAs. Characteristic differences are seen in the NOESY plots of A, B, and Z forms of DNA.[670,671,673,674] Although detailed structural information has been obtained for short segments of DNA, spectra of larger oligonucleotides are impossible to analyze with two-dimensional methods because of extensive overlap of resonances.[665] The difficulty is already apparent in the 17 base pair DNA segment for which a one-dimensional spectrum as well as COSY and NOESY spectra are shown in Fig. 5-53.

Some help with the complexity can be obtained by incorporation of $^{13}C$-enriched methyl groups into

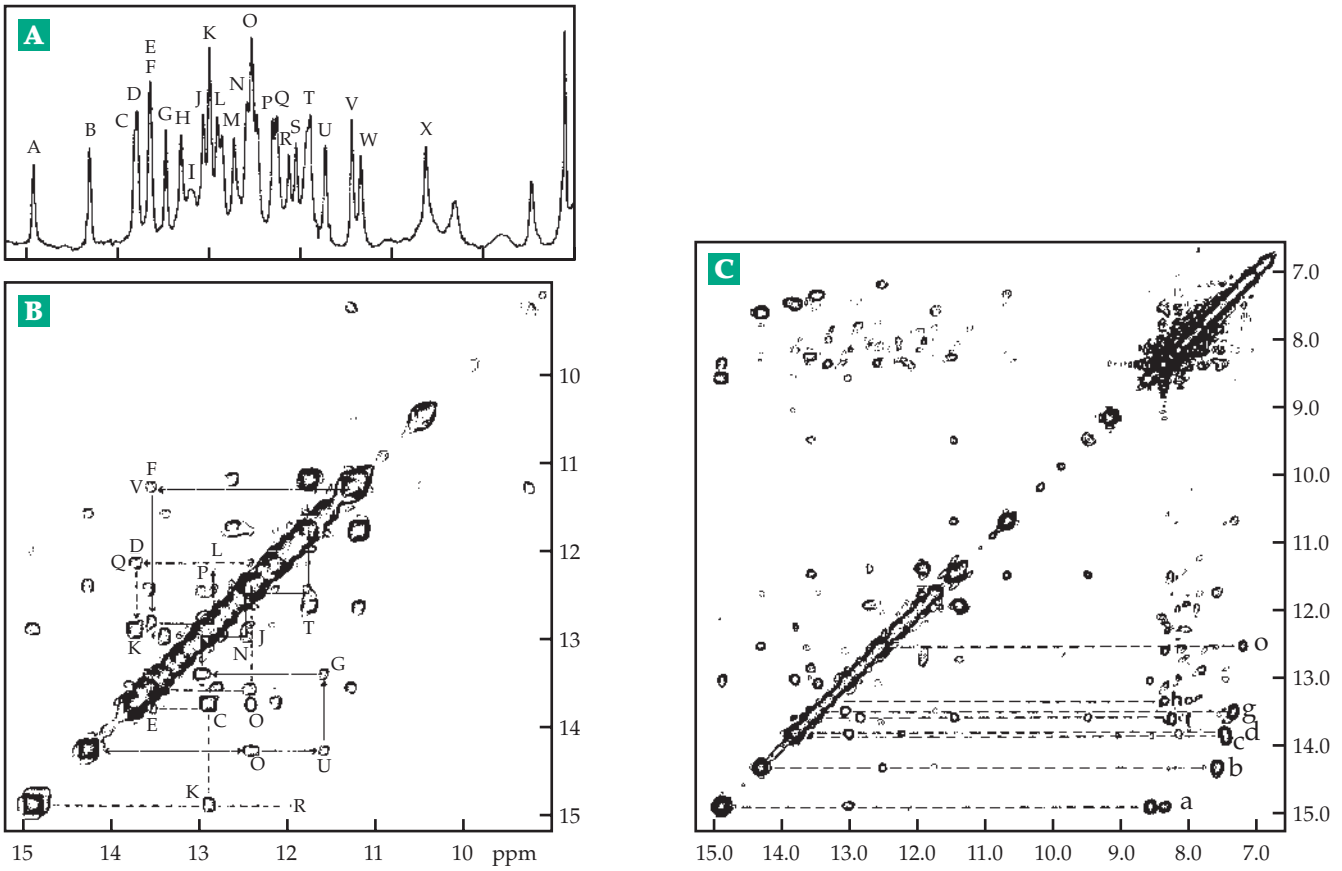



**Figure 5-51** (A) The low-field region of the one-dimensional $^1$H NMR spectrum of *E. coli* $tRNA_1^{Val}$ at 27°C in $H_2O$. Resonances are identified by letters A – X. (B) NOESY spectrum of the same tRNA under similar conditions showing the imino-imino NOEs. In the lower right sector the connectivity traces of the acceptor helix and dihydrouridine helix are shown as solid and dotted lines, respectively. In the NOESY sample the two protons in peak EF are partially resolved whereas the two protons in peak T have coalesced. (C) NOESY spectrum of *E. coli* $tRNA_1^{Val}$ at 32°C showing the imino and aromatic proton regions. AU-type imino protons have been connected horizontally by a dotted line to the cross-peak of their proximal C2-H or C8-H in the 7 to 9 ppm region, which has been labeled with the corresponding lower-case letter. From Hare *et al.*[669] Courtesy of Brian Reid.

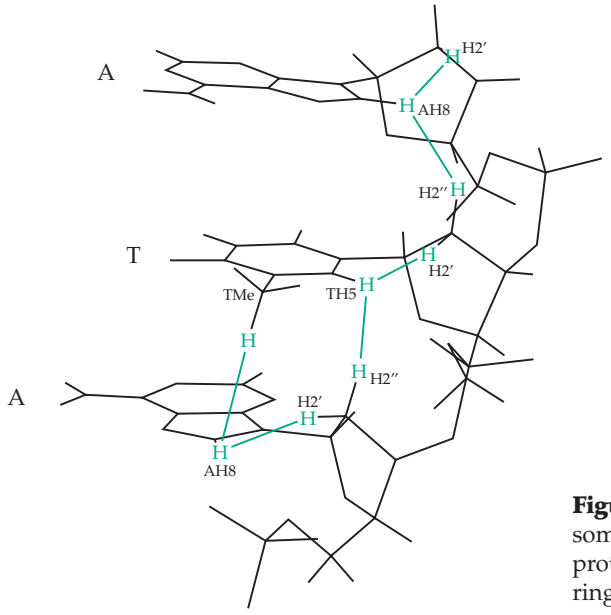



**Figure 5-52** Segment of a DNA chain in the B conformation illustrating some intrachain NOEs that may be observed. The close juxtaposition of proton pairs is provided by the H2'-*endo* ($^2$E) conformation of the sugar rings with *anti* base conformation (Fig. 5-11). After Cohen.[675]

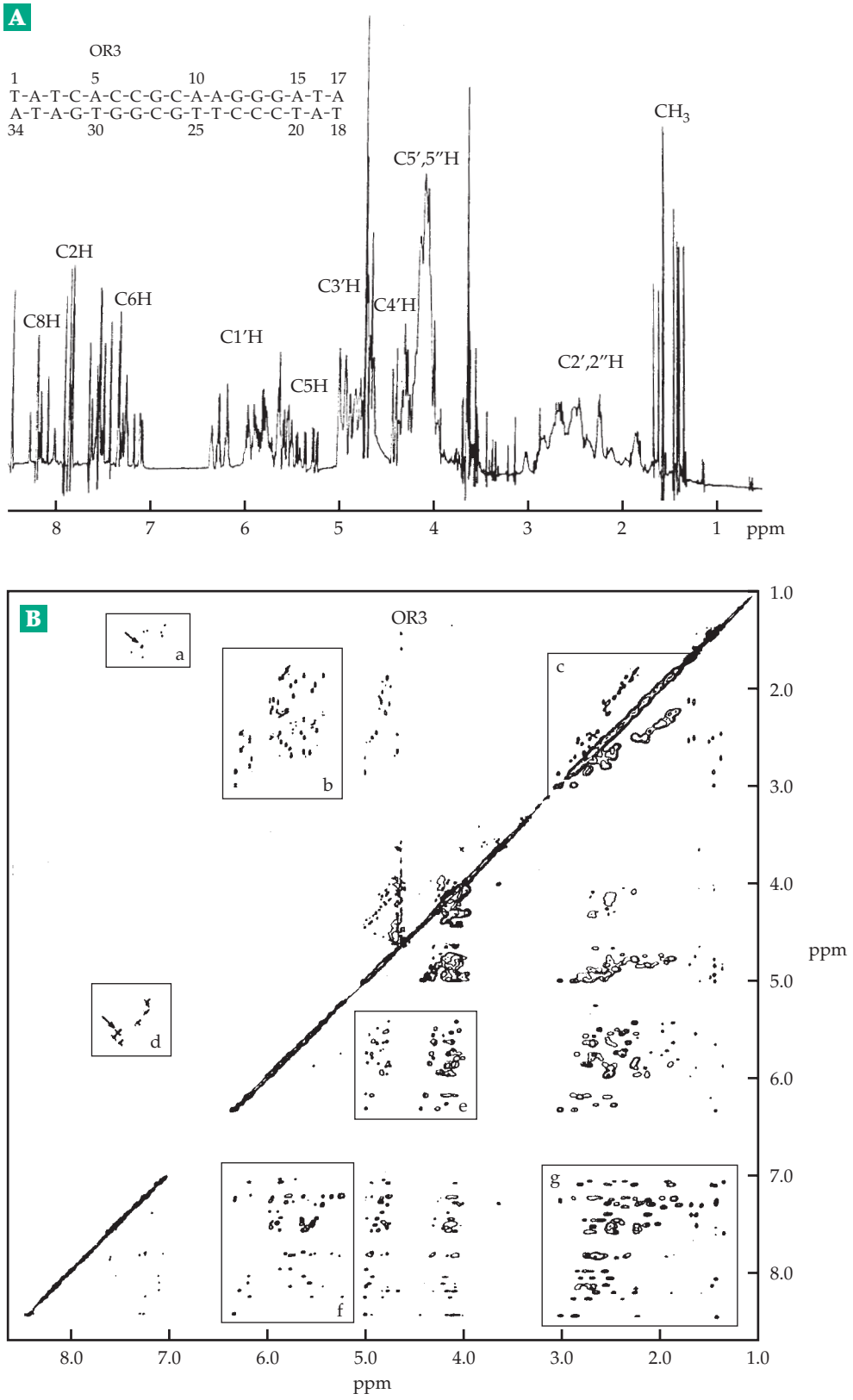



**Figure 5-53** (A) $^1$H NMR spectrum of a 17 base-pair DNA segment from the operator sequence OR3 from bacteriophage λ in $D_2O$ at 37°C. (B) Combined COSY above the diagonal and NOESY (below the diagonal) spectra. C5H and C6H J coupling is established from cross-peaks in box d for cytosines and in box a for thymines. Two unresolved cross-peaks give rise to the more intense spots marked by arrows. Box b contains cross-peaks from scalar coupling of the two H2' protons to the H1' protons of the deoxyribose rings. Most of the aromatic proton resonances could be assigned using the NOE cross-peaks in box f. For further details see Wemmer *et al.*[676] See also Bax and Lerner.[672] Courtesy of B. Reid.

polynucleotides. One way to do this is to grow yeast on a medium containing methionine with $^{13}C$ in its methyl group. The most intense peaks in $^{13}C$ NMR spectra of tRNA molecules from this yeast will represent methyl groups in modified bases in the tRNA. Incorporated $^{13}C$ also permits study of internal motion within tRNA or other oligonucleotides by NMR methods.[677] However, as with protein NMR spectroscopy the major recent advances have come from systematic incorporation of both $^{13}C$ and $^{15}N$ into nucleic acids and the development of three- and four-dimensional NMR methods.[678–684b] Also important are methods for replacing some hydrogen atoms with deuterium to simplify spectra.[685,686]

The very sensitive $^{19}F$ nucleus can be introduced into tRNAs by incorporation of 5-fluorouracil in place of uracil[687] (Fig. 5-54A,B). Phosphorus 31 NMR spectra (Fig. 5-54C) can provide information about conformations of the chain.[688]

A few recent NMR investigations of polynucleotides include studies of triple-helical DNA,[689] Holliday junctions,[290] double-stranded oligonucleotides containing adducts of carcinogens,[690,691] of hairpin loops with sheared A•A and G•G pairs,[692] and of proton exchange in both imino and amino groups.[693]

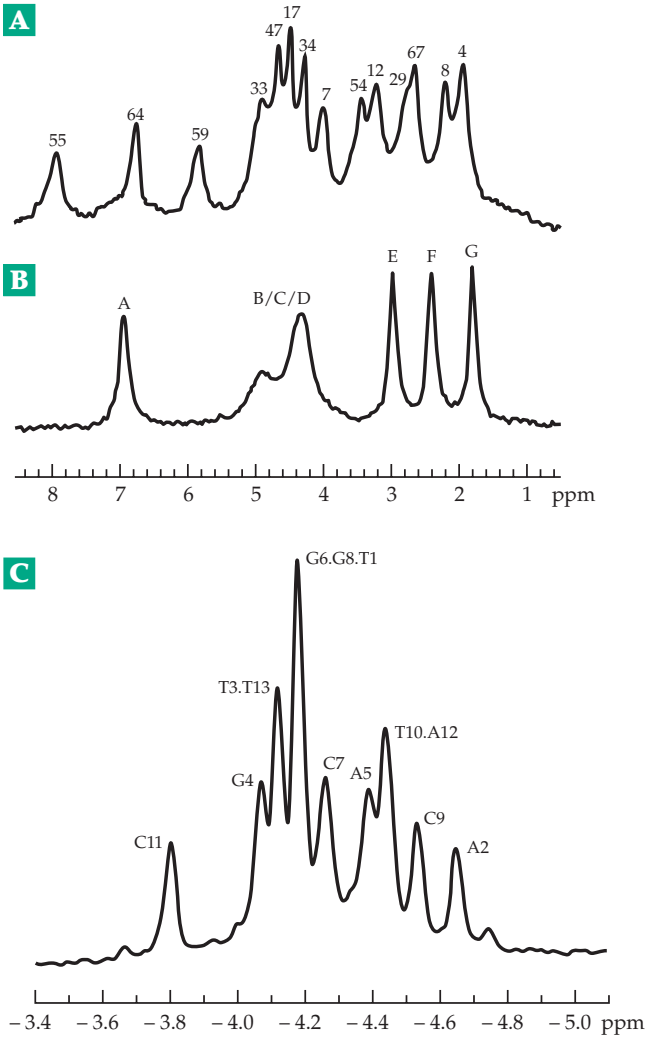



**Figure 5-54** (A) An $^{19}F$ NMR spectrum of the 76-residue *E. coli* $tRNA^{Val}$ containing 5-fluorouracil in 14 positions. Recorded at 47°C. The numbers above the resonances indicate the position in the sequence. (The sequence is not identical to that for the yeast tRNA shown in Fig. 5-30.) Modified from Chu *et al.*[694] Courtesy of Jack Horowitz. (B) A similar spectrum for a 35-residue "minihelix" that contains the acceptor stem of the $tRNA^{Val}$ and seven fluorouracils. The broad peaks B, D, and E are shifted far upfield by reaction with bisulfite (Eq. 5-11) suggesting that they are not hydrogen bonded and are present in the loop of the stem–loop structure. Peaks A, E, F, and G correspond to resonances 64, 7, 67, and 4, respectively, in (A) and represent fluorouracil in the stem structure. From Chu *et al.*[694] Courtesy of Jack Horowitz. (C) A $^{31}P$ NMR spectrum of a synthetic 14 base-pair DNA segment related to the *E. coli lac* operator. The palindromic sequence is TCTGAGCGCTCAGA. The numbers refer to the positions from the 5' end. From Schroeder *et al.*[688]

## References


1. Portugal, F. H., and Cohen, J. S. (1977) *A Century of DNA: A History of the Discovery of the Structure and Function of the Genetic Substance*, MIT Press, Cambridge, Massachusetts
2. Levene, P. A., and Tipson, R. S. (1935) *J. Biol. Chem.* **109**, 623–630
3. Schlenk, F. (1988) *Trends Biochem. Sci.* **13**, 67–69
4. Watson, J. D., and Crick, F. H. C. (1953) *Nature (London)* **171**, 737–738
5. Watson, J. D. (1968) *The Double Helix*, Atheneum, New York
6. Crick, F. (1988) *What a Mad Pursuit: A Personal View of Scientific Discovery*, Basic Books, New York
7. Olby, R. (1974) *The Path to the Double Helix*, Macmillan, London
8. Wilson, H. R. (1988) *Trends Biochem. Sci.* **13**, 275–278
9. Chambers, D. A., Reid, K. B. M., and Cohen, R. L. (1994) *FASEB J.* **8**, 1219–1226

9a. Piper, A. (1998) *Trends Biochem. Sci.* **23**, 151–154

10. Arnott, S., and Hukins, D. W. L. (1973) *J. Mol. Biol.* **81**, 93–105
11. Proudfoot, N. J., Shander, M. H. M., Manley, J. L., Gefter, M. L., and Maniatis, T. (1980) *Science* **209**, 1329–1336
12. Tucker, P. W., Liu, C.-P., Mushinski, J. F., and Blattner, F. R. (1980) *Science* **209**, 1353–1360
13. Yamada, Y., Mudryj, M., Sullivan, M., and de Crombrugghe, B. (1983) *J. Biol. Chem.* **258**, 2758–2761
14. Tsuzuki, T., Obaru, K., Setoyama, C., and Shimada, K. (1987) *J. Mol. Biol.* **198**, 21–31
15. Ratner, L., and 18 other authors (1985) *Nature (London)* **313**, 277–284
16. Dunn, J. J., and Studier, F. W. (1983) *J. Mol. Biol.* **166**, 477– 535
17. Peden, K. W. C. (1983) *Gene* **22**, 277–280

## References


18. Anderson, S., Bankier, A. T., Barrell, B. G., deBruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1981) *Nature (London)* **290**, 457–470
19. Fraser, C. M., and 28 other authors (1995) *Science* **270**, 397–403
20. Fleischmann, R. D., and 39 other authors (1995) *Science* **269**, 496–512

20a. Doolittle, R. F. (1998) *Nature (London)* **392**, 339–342

21. Nowak, R. (1995) *Science* **267**, 172–174
22. Dujon, B., and 107 other authors (1994) *Nature (London)* **369**, 371–378
23. Johnston, M., and 34 other authors (1994) *Science* **265**, 2077–2082
24. Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., and Oliver, S. G. (1996) *Science* **274**, 546–557

24a. Macilwain, C. (2000) *Nature (London)* **405**, 983–984

25. Liébecq, C., ed. (1992) *Biochemical Nomenclature*, Portland Press, London and Chapel Hill, North Carolina (for the International Union of Biochemistry and Molecular Biology)
26. Shabarova, Z., and Bogdanov, A. (1994) *Advanced Organic Chemistry of Nucleic Acids*, VCH Publ., Weinheim
27. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer, New York
28. Hurst, D. T. (1980) *An Introduction to the Chemistry and Biochemistry of Pyrimidines, Purines and Pteridines*, Wiley, New York
29. Markowski, V., Sullivan, G. R., and Roberts, J. D. (1977) *J. Am. Chem. Soc.* **99**, 714–718
30. Katritzky, A. R., and Karelson, M. (1991) *J. Am. Chem. Soc.* **113**, 1561–1566
31. Szczesniak, M., Leszczynski, J., and Person, W. B. (1992) *J. Am. Chem. Soc.* **114**, 2731–2733
32. Scanlan, M. J., and Hillier, I. H. (1984) *J. Am. Chem. Soc.* **106**, 3737–3745
33. Brown, D. M. (1974) in *Basic Principles in Nucleic Acid Chemistry*, Vol. II (Tso, P. O. P., ed), pp. 2–90, Academic Press, New York
34. Hoogsteen, K. (1963) *Acta Crystallogr.* **16**, 907–916
35. Frank-Kamenetskii, M. D., and Mirkin, S. M. (1995) *Ann. Rev. Biochem.* **64**, 65–95
36. Portugal, J. (1989) *Trends Biochem. Sci.* **14**, 127–130
37. Adams, R. L. P., Knowler, J. T., and Leader, D. P. (1992) *The Biochemistry of the Nucleic Acids*, 11th ed., Chapman & Hall, London

37a. Saenger, W. (1973) *Angew. Chem. Int. Ed. Engl.* **12**, 591–601

38. Fang, G., and Cech, T. R. (1993) *Biochemistry* **32**, 11646–11657
39. Gellert, M., Lipsett, M. N., and Davies, D. R. (1962) *Proc. Natl. Acad. Sci. U.S.A.* **48**, 2013–2018
40. Kennell, D. E. (1971) *Prog. Nucleic Acid Res. Mol. Biol.* **11**, 259–301

40a. Stofer, E., Chipot, C., and Lavery, R. (1999) *J. Am. Chem. Soc.* **121**, 9503–9508

40b. Peyret, N., Seneviratne, P. A., Allawi, H. T., and SantaLucia, J., Jr. (1999) *Biochemistry* **38**, 3468–3477

41. Pranata, J., Wierschke, S. G., and Jorgensen, W. L. (1991) *J. Am. Chem. Soc.* **113**, 2810–2819
42. Dang, L. X., and Kollman, P. A. (1990) *J. Am. Chem. Soc.* **112**, 503–507
43. Newcomb, L. F., and Gellman, S. H. (1994) *J. Am. Chem. Soc.* **116**, 4993–4994
44. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) *J. Mol. Biol.* **86**, 843–853
45. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9373–9377
46. Sarai, A., Mazur, J., Nussinov, R., and Jernigan, R. L. (1988) *Biochemistry* **27**, 8498–8502
47. Hunter, C. A. (1993) *J. Mol. Biol.* **230**, 1025–1054
48. Friedman, R. A., and Honig, B. (1995) *Biophys. J.* **69**, 1528–1535
49. DePrisco Albergo, D., Marky, L. A., Breslauer, K. J., and Turner, D. H. (1981) *Biochemistry* **20**, 1409–1413
50. Privalov, P. L., and Filimonov, V. V. (1978) *J. Mol. Biol.* **122**, 447–464
51. Privalov, P. L., and Filimonov, V. V. (1978) *J. Mol. Biol.* **122**, 465–470
52. Petersen, S. B., and Led, J. J. (1981) *J. Am. Chem. Soc.* **103**, 5308–5313
53. Petruska, J., and Goodman, M. F. (1995) *J. Biol. Chem.* **270**, 746–750
54. SantaLucia, J., Jr., Allawi, H. T., and Seneviratne, P. A. (1996) *Biochemistry* **35**, 3555–3562
55. Egli, M., Usman, N., and Rich, A. (1993) *Biochemistry* **32**, 3221–3237
56. Watson, J. D., and Crick, F. H. C. (1953) *Nature (London)* **171**, 964–967
57. Topal, M. D., and Fresco, J. R. (1976) *Nature (London)* **263**, 285–289
58. Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. (1973) *Nature New Biol.* **246**, 40–41
59. Fazakerley, G. V., Gdaniec, Z., and Sowers, L. C. (1993) *J. Mol. Biol.* **230**, 6–10
60. Florián, J., Hrouda, V., and Hobza, P. (1994) *J. Am. Chem. Soc.* **116**, 1457–1460
61. Douhal, A., Kim, S. K., and Zewail, A. H. (1995) *Nature (London)* **378**, 260–263
62. Gabb, H. A., and Harvey, S. C. (1993) *J. Am. Chem. Soc.* **115**, 4218–4227
63. Ellervik, U., and Magnusson, G. (1994) *J. Am. Chem. Soc.* **116**, 2340–2347
64. Thibaudeau, C., Plavec, J., and Chattopadhyaya, J. (1994) *J. Am. Chem. Soc.* **116**, 8033–8037
65. Santos, R. A., Tang, P., and Harbison, G. S. (1989) *Biochemistry* **28**, 9372–9378
66. Sundaralingam, M. (1969) *Biopolymers* **7**, 821–860
67. Vologodskii, A. (1992) *Topology and Physics of Circular DNA*, CRC Press, Boca Raton, Florida
68. Baikalov, I., Grzeskowiak, K., Yanagi, K., Quintana, J., and Dickerson, R. E. (1993) *J. Mol. Biol.* **231**, 768–784
69. Lipanov, A., Kopka, M. L., Kaczor-Grzeskowiak, M., Quintana, J., and Dickerson, R. E. (1993) *Biochemistry* **32**, 1373–1389
70. Schlick, T., Hingerty, B. E., Peskin, C. S., Overton, M. L., and Broyde, S. (1991) *Theoretical Biochemistry and Molecular Physics*, Vol. 1 (Beveridge, D. L. and Lavery, R., eds), pp. 39–58, Adenine Press, Guilderland, New York
71. Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G., and Rich, A. (1979) *Nature (London)* **282**, 680–686
72. Kornberg, A., and Baker, T. A. (1992) *DNA Replication*, 2nd ed., Freeman, New York
73. Kaufmann, W. K., and Paules, R. S. (1996) *FASEB J.* **10**, 238–247
74. Lindahl, T. (1993) *Nature (London)* **362**, 709–715
75. Rodley, G. A., and Bates, R. H. T. (1980) *Trends Biochem. Sci.* **5**, 231–233
76. Kennard, O., and Salisbury, S. A. (1993) *J. Biol. Chem.* **268**, 10701–10704
77. Drew, H. R., and Dickerson, R. E. (1981) *J. Mol. Biol.* **151**, 535–556
78. Dickerson, R. E. (1983) *Sci. Am.* **249**(Dec), 94–111

78a. Dickerson, R. E. (1997) in *International Tables for Macromolecular Crystallography* (Rossmann, M. G., and Arnold, E., eds), (Chapter 23, Part c)

79. Shui, X., McFail-Isom, L., Hu, G. G., and Williams, L. D. (1998) *Biochemistry* **37**, 8341–8355
80. Gorin, A. A., Zhurkin, V. B., and Olson, W. K. (1995) *J. Mol. Biol.* **247**, 34–48
81. Tereshko, V., Minasov, G., and Egli, M. (1999) *J. Am. Chem. Soc.* **121**, 470–471
82. Yuan, H., Quintana, J., and Dickerson, R. E. (1992) *Biochemistry* **31**, 8009–8021
83. Ban, C., Ramakrishnan, B., and Sundaralingam, M. (1994) *J. Mol. Biol.* **236**, 275–285

83a. Kielkopf, C. L., Ding, S., Kuhn, P., and Rees, D. C. (2000) *J. Mol. Biol.* **296**, 787–801

84. Johansson, E., Parkinson, G., and Neidle, S. (2000) *J. Mol. Biol.* **300**, 551–561
85. Shen, L. X., Cai, Z., and Tinoco, I., Jr. (1995) *FASEB J.* **9**, 1023–1033
86. Rosenberg, J. M., Seeman, N. C., Day, R. O., and Rich, A. (1976) *J. Mol. Biol.* **104**, 145–167
87. Marino, J. P., Gregorian, R. S., Jr., Csankovszki, G., and Crothers, D. M. (1995) *Science* **268**, 1448–1454
88. Iwamoto, S., and Hsu, M.-T. (1983) *Nature (London)* **305**, 70–72
89. Morel, G., ed. (1995) *Visualization of Nucleic Acids*, CRC Press, Boca Raton, Florida
90. Arscott, P. G., Lee, G., Bloomfield, V. A., and Evans, D. F. (1989) *Nature (London)* **339**, 484–486
91. Jares-Erijman, E. A., and Jovin, T. M. (1996) *J. Mol. Biol.* **257**, 597–617
92. Duan, Y., Wilkosz, P., Crowley, M., and Rosenberg, J. M. (1997) *J. Mol. Biol.* **272**, 553–572
93. Auffinger, P., and Westhof, E. (2000) *J. Mol. Biol.* **300**, 1113–1131

93a. Soler-López, M., Malinina, L., Liu, J., Huynh-Dinh, T., and Subirana, J. A. (1999) *J. Biol. Chem.* **274**, 23683–23686

94. Schneider, B., and Berman, H. M. (1995) *Biophys. J.* **69**, 2661–2669
95. Tereshko, V., Minasov, G., and Egli, M. (1999) *J. Am. Chem. Soc.* **121**, 3590–3595
96. Franklin, R. E., and Gosling, R. G. (1953) *Nature (London)* **172**, 156–157
97. Ramakrishnan, B., and Sundaralingam, M. (1995) *Biophys. J.* **69**, 553–558
98. Ramakrishnan, B., and Sundaralingam, M. (1993) *J. Mol. Biol.* **231**, 431–444
99. Eisenstein, M., and Shakked, Z. (1995) *J. Mol. Biol.* **248**, 662–678

99a. Ng, H.-L., Kopka, M. L., and Dickerson, R. E. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2035–2039

100. Selsing, E., Wells, R. D., Alden, C. J., and Arnott, S. (1979) *J. Biol. Chem.* **254**, 5417–5422
101. Mahendrasingam, A., Forsyth, V. T., Hussain, R., Greenall, R. J., Pigram, W. J., and Fuller, W. (1986) *Science* **233**, 195–197
102. Hopkins, R. C. (1981) *Science* **211**, 289–291
103. Gupta, G., Bansal, M., and Sasisekharan, V. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6486–6490
104. Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A. G. W., and Ratliff, R. L. (1980) *Nature (London)* **283**, 743–745
105. Herbert, A., and Rich, A. (1996) *J. Biol. Chem.* **271**, 11595–11598
106. Gessner, R. V., Frederick, C. A., Quigley, G. J., Rich, A., and Wang, A. H.-J. (1989) *J. Biol. Chem.* **264**, 7921–7935
107. Zhou, G., and Ho, P. S. (1990) *Biochemistry* **29**, 7229–7236

## References


108. Parkinson, G. N., Arvanitis, G. M., Lessinger, L., Ginell, S. L., Jones, R., Gaffney, B., and Berman, H. M. (1995) *Biochemistry* **34**, 15487–15495
109. Sadasivan, C., and Gautham, N. (1995) *J. Mol. Biol.* **248**, 918– 930
110. Bancroft, D., Williams, L. D., Rich, A., and Egli, M. (1994) *Biochemistry* **33**, 1073–1086
111. Rich, A., Nordheim, A., and Wang, A. H.-J. (1984) *Ann. Rev. Biochem.* **53**, 791–846
112. Moore, M. H., Van Meervelt, L., Salisbury, S. A., Lin, P. K. T., and Brown, D. M. (1995) *J. Mol. Biol.* **251**, 665–673
113. Wang, Y., Thomas, G. A., and Peticolas, W. L. (1987) *Biochemistry* **26**, 5178–5186
114. Saenger, W., Hunter, W. N., and Kennard, O. (1986) *Nature (London)* **324**, 385–388
115. Frank-Kamenetskii, M. (1986) *Nature (London)* **324**, 305
116. Basham, B., Schroth, G. P., and Ho, P. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6464–6468
117. Misra, V. K., and Honig, B. (1996) *Biochemistry* **35**, 1115–1124
118. McLean, M. J., Blaho, J. A., Kilpatrick, M. W., and Wells, R. D. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5884–5888
119. Kim, J.-m, and DasSarma, S. (1996) *J. Biol. Chem.* **271**, 19724–19731
120. Nordheim, A., Pardue, M. L., Lafer, E. M., Möller, A., Stollar, B. D., and Rich, A. (1981) *Nature (London)* **294**, 417–422
121. Jaworski, A., Hsieh, W.-T., Blaho, J. A., Larson, J. E., and Wells, R. D. (1987) *Science* **238**, 773–777
122. Zacharias, W., O'Conner, T. R., and Larson, J. E. (1988) *Biochemistry* **27**, 2970–2978
123. Nordheim, A., and Rich, A. (1983) *Nature (London)* **303**, 674–679
124. Wells, R. D. (1988) *J. Biol. Chem.* **263**, 1095–1098
125. Rentzeperis, D., Kupke, D. W., and Marky, L. A. (1994) *Biochemistry* **33**, 9588–9591
126. Liu, K., Miles, H. T., Frazier, J., and Sasisekharan, V. (1993) *Biochemistry* **32**, 11802–11809
127. Zhou, N., Germann, M. W., van de Sande, J. H., Pattabiraman, N., and Vogel, H. J. (1993) *Biochemistry* **32**, 646–656
128. Otto, C., Thomas, G. A., Rippe, K., Jovin, T. M., and Peticolas, W. L. (1991) *Biochemistry* **30**, 3062–3069
129. Trifonov, E. N. (1991) *Trends Biochem. Sci.* **16**, 467–470
130. Rich, A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13999–14000
131. Hageman, P. J. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 265–286
132. Poncin, M., Hartmann, B., and Lavery, R. (1992) *J. Mol. Biol.* **226**, 775–794
133. MacKerell, A. D., Jr., Wiórkiewicz-Kuczera, J., and Karplus, M. (1995) *J. Am. Chem. Soc.* **117**, 11946–11975
134. Kosikov, K. M., Gorin, A. A., Zhurkin, V. B., and Olson, W. K. (1999) *J. Mol. Biol.* **289**, 1301–1326
135. Peticolas, W. L., Wang, Y., and Thomas, G. A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2579–2583
136. Dickerson, R. E., and Drew, H. R. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7318–7322

136a. Packer, M. J., and Hunter, C. A. (1998) *J. Mol. Biol.* **280**, 407–420

137. Hogan, M., Dattagupta, N., and Crothers, D. M. (1979) *Nature (London)* **278**, 521–524
138. Mitra, C. K., Sarma, M. H., and Sarma, R. H. (1981) *J. Am. Chem. Soc.* **103**, 6727–6737
139. Cluzel, P., Lebrun, A., Heller, C., Lavery, R., Viovy, J.-L., Chatenay, D., and Caron, F. (1996) *Science* **271**, 792–794
140. Smith, S. B., Cui, Y., and Bustamante, C. (1996) *Science* **271**, 795–799
141. Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A., and Croquette, V. (1996) *Science* **271**, 1835–1837

141a. Allemand, J. F., Bensimon, D., Lavery, R., and Croquette, V. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14152–14157

142. Dickerson, R. E. (1983) *J. Mol. Biol.* **166**, 419–441
143. Tung, C.-S., and Harvey, S. C. (1986) *J. Biol. Chem.* **261**, 3700–3709
144. Ulanovsky, L. E., and Trifonov, E. N. (1987) *Nature (London)* **326**, 720–722
145. Diekmann, S. (1989) *EMBO J.* **8**, 1–4
146. El Hassan, M. A., and Calladine, C. R. (1996) *J. Mol. Biol.* **259**, 95–103
147. Calladine, C. R. (1982) *J. Mol. Biol.* **161**, 343–352
148. Hunter, C. A., and Lu, X.-J. (1997) *J. Mol. Biol.* **265**, 603–619
149. Babcock, M. S., Pednault, E. P. D., and Olson, W. K. (1994) *J. Mol. Biol.* **273**, 125–156
150. Lu, X.-J., El Hassan, M. A., and Hunter, C. A. (1997) *J. Mol. Biol.* **273**, 668–680
151. Kitchin, P. A., Klein, V. A., Ryan, K. A., Gann, K. L., Rauch, C. A., Kang, D. S., Wells, R. D., and England, P. T. (1986) *J. Biol. Chem.* **261**, 11302–11309
152. Koo, H.-S., Wu, H.-M., and Crothers, D. M. (1986) *Nature (London)* **320**, 501–506
153. Zahn, K., and Blattner, F. B. (1987) *Science* **236**, 416–422
154. Snyder, M., Buchman, A. R., and Davis, R. W. (1986) *Nature (London)* **324**, 87–89
155. Travers, A. A. (1987) *Trends Biochem. Sci.* **12**, 108–112
156. Grzeskowiak, K., Goodsell, D. S., Kaczor-Grzeskowiak, M., Cascio, D., and Dickerson, R. E. (1993) *Biochemistry* **32**, 8923–8931
157. Allewell, N. (1988) *Trends Biochem. Sci.* **13**, 193–195
158. Haran, T. E., Kahn, J. D., and Crothers, D. M. (1994) *J. Mol. Biol.* **244**, 135–143
159. Strauss, J. K., and Maher, L. J., III. (1994) *Science* **266**, 1829–1834
160. Olson, W. K., Marky, N. L., Jernigan, R. L., and Zhurkin, V. B. (1993) *J. Mol. Biol.* **232**, 530–554
161. Price, M. A., and Tullius, T. D. (1993) *Biochemistry* **32**, 127–136
162. DiGabriele, A. D., and Steitz, T. A. (1993) *J. Mol. Biol.* **231**, 1024–1039
163. Chuprina, V. P., Fedoroff, O. Y., and Reid, B. R. (1991) *Biochemistry* **30**, 561–568
164. Crothers, D. M. (1994) *Science* **266**, 1819–1820
165. Husain, I., Griffith, J., and Sancar, A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2558–2562
166. Srinivasan, A. R., Torres, R., Clark, W., and Olson, W. K. (1987) *J. Biomol. Struct. Dyn.* **5**, 459–496
167. Lutter, L. C., Halvorson, H. R., and Calladine, C. R. (1996) *J. Mol. Biol.* **261**, 620–633
168. Privé, G. G., Heinemann, V., Chandrasegaran, S., Kan, L.-S., Kopka, M. L., and Dickerson, R. E. (1987) *Science* **238**, 498–504
169. Lilley, D. M. J. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7140–7142
170. Woodson, S. A., and Crothers, D. M. (1988) *Biochemistry* **27**, 3130–3141
171. Miller, M., Harrison, R. W., Wlodawer, A., Appella, E., and Sussman, J. L. (1988) *Nature (London)* **334**, 85–86
172. Hunter, W. N., Brown, T., Kneale, G., Anand, N. N., Rabinovich, D., and Kennard, O. (1987) *J. Biol. Chem.* **262**, 9962–9970
173. Lane, A., Martin, S. R., Ebel, S., and Brown, T. (1992) *Biochemistry* **31**, 12087–12095
174. Li, Y., and Agrawal, S. (1995) *Biochemistry* **34**, 10056–10062
175. Manning, G. S. (1978) *Q. Rev. Biophys.* **11**, 179–246
176. Mascotti, D. P., and Lohman, T. M. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3142–3146
177. Li, A. Z., Huang, H., Re, X., Qi, L. J., and Marx, K. A. (1998) *Biophys. J.* **74**, 964–973
178. Frydman, L., Rossomando, P. C., Frydman, V., Fernandez, C. O., Frydman, B., and Samejima, K. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9186–9190
179. Jain, S., Zon, G., and Sundaralingam, M. (1989) *Biochemistry* **28**, 2360–2364
180. Pelta, J., Durand, D., Doucet, J., and Livolant, F. (1996) *Biophys. J.* **71**, 48–63
181. Rajski, S. R., Kumar, S., Roberts, R. J., and Barton, J. K. (1999) *J. Am. Chem. Soc.* **121**, 5615–5616

181a. Aich, P., Labiuk, S. L., Tari, L. W., Dalbaere, L. J. T., Roesler, W. J., Falk, K. J., Steer, R. P., and Lee, J. S. (1999) *J. Mol. Biol.* **294**, 477–485

182. Fink, H.-W., and Schönenberger, C. (1999) *Nature (London)* **398**, 407–410
183. Marini, J. C., Miller, K. G., and Englund, P. T. (1980) *J. Biol. Chem.* **255**, 4976–4979
184. Wasserman, S. A., Dungan, J. M., and Cozzarelli, N. R. (1985) *Science* **229**, 171–174
185. Du, S. M., Wang, H., Tse-Dinh, Y.-C., and Seeman, N. C. (1995) *Biochemistry* **34**, 673–682
186. Du, S. M., Stollar, B. D., and Seeman, N. C. (1995) *J. Am. Chem. Soc.* **117**, 1194–1200
187. Wasserman, S. A., and Cozzarelli, N. R. (1986) *Science* **232**, 951–960
188. White, J. H., and Cozzarelli, N. R. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3322–3326
189. Cipra, B. (1992) *Science* **255**, 403
190. Adams, C. C. (1994) *The Knot Book*, Freeman, New York
191. Krasnow, M. A., Stasiak, A., Spengler, S. J., Dean, F., Koller, T., and Cozzarelli, N. R. (1983) *Nature (London)* **304**, 559–560
192. Du, S. M., and Seeman, N. C. (1992) *J. Am. Chem. Soc.* **114**, 9652–9655
193. Wang, J. C. (1980) *Trends Biochem. Sci.* **5**, 219–221
194. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry, Part 3*, Freeman, San Francisco, California (pp. 1265–1290)
195. Lebowitz, J. (1990) *Trends Biochem. Sci.* **15**, 202–207
196. Bates, A. D., and Maxwell, A. (1993) *DNA Topology*, Oxford Univ. Press, London
197. Yang, Y., Westcott, T. P., Pedersen, S. C., Tobias, I., and Olson, W. K. (1995) *Trends Biochem. Sci.* **20**, 313–319
198. Vinograd, J., Lebowitz, J., and Watson, R. (1968) *J. Mol. Biol.* **33**, 173–197
199. White, J. H., Cozzarelli, N. R., and Bauer, W. R. (1988) *Science* **241**, 323–327

199a. Vologodskii, A. V., and Cozzarelli, N. R. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 609–643

199b. Cozzarelli, N. R., Boles, T. C., and White, J. H. (1990) in *DNA Topology and its Biological Effects*, pp. 139–184, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York

199c. Podtelezhnikov, A. A., Cozzarelli, N. R., and Vologodskii, A. V. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12974–12979

200. Reich, Z., Levin-Zaidman, S., Gutman, S. B., Arad, T., and Minsky, A. (1994) *Biochemistry* **33**, 14177–14184
201. Eriksson, M. A. L., Härd, T., and Nilsson, L. (1995) *Biophys. J.* **68**, 402–426
202. Keller, W. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2550–2554
203. Kornberg, R. D., and Klug, A. (1981) *Sci. Am.* **244**(Feb), 52–64
204. Wang, B.-C., Rose, J., Arents, G., and Moudrianakis, E. N. (1994) *J. Mol. Biol.* **236**, 179–188

## References


205. Malik, M., Bensaid, A., Rouviere-Yaniv, J., and Drlica, K. (1996) *J. Mol. Biol.* **256**, 66–76
206. Serrano, M., Salas, M., and Hermoso, J. M. (1993) *Trends Biochem. Sci.* **18**, 202–206
207. Seidl, A., and Hinz, H.-J. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1312–1316
208. Rahmouni, A. R., and Wells, R. D. (1989) *Science* **246**, 358–363
209. Spolar, R. S., and Record, M. T., Jr. (1994) *Science* **263**, 777–784
210. White, J. H., Gallo, R. M., and Bauer, W. R. (1992) *Trends Biochem. Sci.* **17**, 7–12
211. van Holde, K., and Zlatanova, J. (1995) *J. Biol. Chem.* **270**, 8373–8376
212. Keller, W. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4876–4880
213. Hamiche, A., Schultz, P., Ramakrishnan, V., Oudet, P., and Prunell, A. (1996) *J. Mol. Biol.* **257**, 30–42
214. Alfonso, P. J., Crippa, M. P., Hayes, J. J., and Bustin, M. (1994) *J. Mol. Biol.* **236**, 189–198
215. Bednar, J., Furrer, P., Stasiak, A., Dubochet, J., Egelman, E. H., and Bates, A. D. (1994) *J. Mol. Biol.* **235**, 825–847
216. Kumar, A. (1995) *Biochemistry* **34**, 12921–12925
217. Stigter, D. (1995) *Biophys. J.* **69**, 380–388
218. Baeza, I., Gariglio, P., Rangel, L. M., Chavez, P., Cervantes, L., Arguello, C., Wong, C., and Montañez, C. (1987) *Biochemistry* **26**, 6387–6392
219. Schmid, M. B. (1988) *Trends Biochem. Sci.* **13**, 131–135
220. Quigley, G. J., Wang, A. H.-J., Ughetto, G., van der Marel, G., van Boom, J. H., and Rich, A. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 7204–7208
221. Campbell, A. M., and Jolly, D. J. (1973) *Biochem. J.* **133**, 209–226
222. Depew, R. E., and Wang, J. C. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4275–4279
223. Westof, E., and Sundaralingam, M. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1852–1856
224. Sobell, H. M. (1974) *Sci. Am.* **231**(Aug), 82–91
225. Chen, H., and Patel, D. J. (1995) *J. Am. Chem. Soc.* **117**, 5901–5913
226. Chaires, J. B., Satyanarayana, S., Suh, D., Fokt, I., Przewloka, T., and Priebe, W. (1996) *Biochemistry* **35**, 2047–2053
227. Rajeswari, M. R., Bose, H. S., Kukreti, S., Gupta, A., Chauhan, V. S., and Roy, K. B. (1992) *Biochemistry* **31**, 6237–6241

227a. Leitner, D., Schröder, W., and Weisz, K. (2000) *Biochemistry* **39**, 5886–5892

228. Felsenfeld, G., Davies, D., and Rich, A. (1957) *J. Am. Chem. Soc.* **79**, 2023–2024
229. Felsenfeld, G., and Miles, H. T. (1967) *Ann. Rev. Biochem.* **36**, 407–448
230. Arnott, S., and Selsing, E. (1974) *J. Mol. Biol.* **88**, 509–521
231. Wang, E., Koshlap, K. M., Gillespie, P., Dervan, P. B., and Feigon, J. (1996) *J. Mol. Biol.* **257**, 1052–1069
232. Radhakrishnan, I., and Patel, D. J. (1994) *Biochemistry* **33**, 11405–11416
233. Radhakrishnan, I., and Patel, D. J. (1994) *J. Mol. Biol.* **241**, 600–619
234. Rhee, S., Han, Z.-j, Liu, K., Miles, H. T., and Davies, D. R. (1999) *Biochemistry* **38**, 16810–16815
235. Roberts, R. W., and Crothers, D. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4320–4325
236. van Dongen, M. J. P., Heus, H. A., Wymenga, S. S., van der Marel, G. A., van Boom, J. H., and Hilbers, C. W. (1996) *Biochemistry* **35**, 1733–1739
237. Klysik, J. (1995) *J. Mol. Biol.* **245**, 499–507
238. Srinivasan, A. R., and Olson, W. K. (1998) *J. Am. Chem. Soc.* **120**, 484–491
239. Hanvey, J. C., Shimizu, M., and Wells, R. D. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6292–6296
240. Htun, H., and Dahlberg, J. E. (1989) *Science* **243**, 1571–1576
241. Mirkin, S. M., and Frank-Kamenetskii, M. D. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 541–576
242. Singh, S., Patel, P. K., and Hosur, R. V. (1997) *Biochemistry* **36**, 13214–13222
243. Van Meervelt, L., Vlieghe, D., Dautant, A., Gallois, B., Précigoux, G., and Kennard, O. (1995) *Nature (London)* **374**, 742–744
244. Reddy, G., Jwang, B., Rao, B. J., and Radding, C. M. (1994) *Biochemistry* **33**, 11486–11492
245. Kim, M. G., Zhurkin, V. B., Jernigan, R. L., and Camerini-Otero, R. D. (1995) *J. Mol. Biol.* **247**, 874–889
246. Cox, M. M. (1995) *J. Biol. Chem.* **270**, 26021–26024
247. Jain, S. K., Cox, M. M., and Inman, R. B. (1995) *J. Biol. Chem.* **270**, 4943–4949
248. Best, G. C., and Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117**, 1187–1193
249. Greenberg, W. A., and Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117**, 5016–5022
250. Hashem, G. M., Pham, L. P., Vaughan, M. R., and Gray, D. M. (1998) *Biochemistry* **37**, 61–72
251. Gryaznov, S. M., Lloyd, D. H., Chen, J.-K., Schultz, R. G., DeDionisio, L. A., Ratmeyer, L., and Wilson, W. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5798–5802
252. Betts, L., Josey, J. A., Veal, J. M., and Jordan, S. R. (1995) *Science* **270**, 1838–1841
253. Tomac, S., Sarkar, M., Ratilainen, T., Wittung, P., Nielsen, P. E., Nordén, B., and Gräslund, A. (1996) *J. Am. Chem. Soc.* **118**, 5544–5552

253a. Ray, A., and Nordén, B. (2000) *FASEB J.* **14**, 1041–1060

253b. Izvolsky, K. I., Demidov, V. V., Nielsen, P. E., and Frank-Kamenetskii, M. D. (2000) *Biochemistry* **39**, 10908–10913

254. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., and Dervan, P. B. (1998) *Nature (London)* **391**, 468–471
255. Marsh, T. C., and Henderson, E. (1994) *Biochemistry* **33**, 10718–10724
256. Blackburn, E. H. (1991) *Nature (London)* **350**, 569–573
257. Blackburn, E. H. (1991) *Trends Biochem. Sci.* **16**, 378–381
258. Moyzis, R. K. (1991) *Sci. Am.* **265**(Aug), 48–55
259. Henderson, E., Hardin, C. C., Wolk, S. K., Tinoco, I., Jr., and Blackburn, E. (1987) *Cell* **51**, 899–908
260. Laughlan, G., Murchie, A. I. H., Norman, D. G., Moore, M. H., Moody, P. C. E., Lilley, D. M. J., and Luisi, B. (1994) *Science* **265**, 520–527
261. Kang, C. H., Zhang, X., Ratliff, R., Moyzis, R., and Rich, A. (1992) *Nature (London)* **356**, 126–131
262. Smith, F. W., and Feigon, J. (1993) *Biochemistry* **32**, 8682–8692
263. Balagurumoorthy, P., and Brahmachari, S. K. (1994) *J. Biol. Chem.* **269**, 21858–21869
264. Williamson, J. R. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 703–730
265. Frank-Kamenetskii, M. (1992) *Nature (London)* **356**, 105
266. Wang, K. Y., Swaminathan, S., and Bolton, P. H. (1994) *Biochemistry* **33**, 7517–7527
267. Wang, Y., and Patel, D. J. (1995) *J. Mol. Biol.* **251**, 76–94
268. Laporte, L., and Thomas, G. J., Jr. (1998) *Biochemistry* **37**, 1327–1335
269. Ross, W. S., and Hardin, C. C. (1994) *J. Am. Chem. Soc.* **116**, 6070–6080
270. Sen, D., and Gilbert, W. (1990) *Nature (London)* **344**, 410–414
271. Miura, T., Benevides, J. M., and Thomas, G. J., Jr. (1995) *J. Mol. Biol.* **248**, 233–238
272. Schierer, T., and Henderson, E. (1994) *Biochemistry* **33**, 2240–2246
273. Hardin, C. C., Corregan, M., Brown, B. A., II, and Frederick, L. N. (1993) *Biochemistry* **32**, 5870–5880
274. Cheong, C., and Moore, P. B. (1992) *Biochemistry* **31**, 8406–8414
275. Chen, F.-M. (1995) *J. Biol. Chem.* **270**, 23090–23096
276. Kettani, A., Kumar, R. A., and Patel, D. J. (1995) *J. Mol. Biol.* **254**, 638–656
277. Kang, C. H., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R., and Rich, A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11636–11640
278. Gehring, K., Leroy, J.-L., and Guéron, M. (1993) *Nature (London)* **363**, 561–565
279. Lacroix, L., Mergny, J.-L., Leroy, J.-L., and Hélène, C. (1996) *Biochemistry* **35**, 8715–8722
280. Chen, L., Cai, L., Zhang, X., and Rich, A. (1994) *Biochemistry* **33**, 13540–13546
281. Kang, C. H., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R., and Rich, A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3874–3878
282. Benevides, J. M., Kang, C., and Thomas, G. J., Jr. (1996) *Biochemistry* **35**, 5747–5755
283. Gallego, J., Chou, S.-H., and Reid, B. R. (1997) *J. Mol. Biol.* **273**, 840–856
284. Selsing, E., Wells, R. D., Alden, C. J., and Arnott, S. (1979) *J. Biol. Chem.* **254**, 5417–5422
285. Nishizaki, T., Iwai, S., Ohkubo, T., Kojima, C., Nakamura, H., Kyogoku, Y., and Ohtsuka, E. (1996) *Biochemistry* **35**, 4016–4025
286. Salazar, M., Champoux, J. J., and Reid, B. R. (1993) *Biochemistry* **32**, 739–744
287. Seeman, N. C., and Kallenbach, N. R. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 53–86
288. Zhang, S., Fu, T.-J., and Seeman, N. C. (1993) *Biochemistry* **32**, 8062–8067
289. Pikkemaat, J. A., van den Elst, H., van Boom, J. H., and Altona, C. (1994) *Biochemistry* **33**, 14896–14907
290. Carlström, G., and Chazin, W. J. (1996) *Biochemistry* **35**, 3534–3544
291. Holliday, R. (1964) *Genet. Res. Camb.* **5**, 282–304
292. Bennett, R. J., and West, S. C. (1995) *J. Mol. Biol.* **252**, 213–226
293. Grainger, R. J., Murchie, A. I. H., and Lilley, D. M. J. (1998) *Biochemistry* **37**, 23–32
294. Lilley, D. M. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9513–9515
295. Panyutin, I. G., Biswas, I., and Hsieh, P. (1995) *EMBO J.* **14**, 1819–1826
296. Chen, J., and Seeman, N. C. (1991) *Nature (London)* **350**, 631–633
297. Zhang, Y., and Seeman, N. C. (1994) *J. Am. Chem. Soc.* **116**, 1661–1669

297a. Liu, F., Sha, R., and Seeman, N. C. (1999) *J. Am. Chem. Soc.* **121**, 917–922

297b. Fahlman, R. P., and Sen, D. (1999) *J. Am. Chem. Soc.* **121**, 11079–11085

298. Feagin, J. E. (1990) *J. Biol. Chem.* **265**, 19373–19376
299. Stuart, K. (1991) *Trends Biochem. Sci.* **16**, 68–72
300. Simons, R. W., and Grunberg-Manago, M., eds. (1997) *RNA Structure and Function*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
301. Nicoloso, M., Qu, L.-H., Michot, B., and Bachellerie, J.-P. (1996) *J. Mol. Biol.* **260**, 178–195
302. Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. (1995) *Ann. Rev. Biochem.* **64**, 763–797
303. Draper, D. E. (1996) *Trends Biochem. Sci.* **21**, 145–149
304. Sich, C., Ohlenschläger, O., Ramachandran, R., Görlach, M., and Brown, L. R. (1997) *Biochemistry* **36**, 13989–14002

## References


305. Egli, M., Portmann, S., and Usman, N. (1996) *Biochemistry* **35**, 8489–8494
306. Auffinger, P., and Westhof, E. (1997) *J. Mol. Biol.* **274**, 54–63
307. Salazar, M., Fedoroff, O. Y., Miller, J. M., Ribeiro, N. S., and Reid, B. R. (1993) *Biochemistry* **32**, 4207–4215
308. Fedoroff, O. Y., Salazar, M., and Reid, B. R. (1993) *J. Mol. Biol.* **233**, 509–523
309. Ratmeyer, L., Vinayak, R., Zhong, Y. Y., Zon, G., and Wilson, W. D. (1994) *Biochemistry* **33**, 5298–5304
310. Varani, G., Cheong, C., and Tinoco, I., Jr. (1991) *Biochemistry* **30**, 3280–3289
311. Butcher, S. E., Dieckmann, T., and Feigon, J. (1997) *J. Mol. Biol.* **268**, 348–358
312. van Dongen, M. J. P., Wijmenga, S. S., van der Marel, G. A., van Boom, J. H., and Hilbers, C. W. (1996) *J. Mol. Biol.* **263**, 715–729
313. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1992) *Science* **256**, 217–219
314. Murphy, F. L., and Cech, T. R. (1994) *J. Mol. Biol.* **236**, 49–63
315. Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H. M., and Pardi, A. (1996) *J. Mol. Biol.* **264**, 968–980

315a. Gutell, R. R., Cannone, J. J., Konings, D., and Gautheret, D. (2000) *J. Mol. Biol.* **300**, 791–803

316. Fountain, M. A., Serra, M. J., Krugh, T. R., and Turner, D. H. (1996) *Biochemistry* **35**, 6539–6548
317. Huang, S., Wang, Y.-X., and Draper, D. E. (1996) *J. Mol. Biol.* **258**, 308–321
318. Nowakowski, J., and Tinoco, I., Jr. (1996) *Biochemistry* **35**, 2577–2585
319. Chang, K.-Y., and Tinoco, I., Jr. (1997) *J. Mol. Biol.* **269**, 52–66
320. Rosendahl, G., Hansen, L. H., and Douthwaite, S. (1995) *J. Mol. Biol.* **249**, 59–68
321. Pleij, C. W. A. (1990) *Trends Biochem. Sci.* **15**, 143–147
322. ten Dam, E., Pleij, K., and Draper, D. (1992) *Biochemistry* **31**, 11665–11676
323. Shen, L. X., and Tinoco, I., Jr. (1995) *J. Mol. Biol.* **247**, 963–978
324. Pinard, R., Payant, C., and Brakier-Gingras, L. (1995) *Biochemistry* **34**, 9611–9616
325. Puglisi, J. D., Wyatt, J. R., and Tinoco, I., Jr. (1988) *Nature (London)* **331**, 283–286
326. Qiu, H., Kaluarachchi, K., Du, Z., Hoffman, D. W., and Giedroc, D. P. (1996) *Biochemistry* **35**, 4176–4186

326a. Kolk, M. H., van der Graaf, M., Wijmenga, S. S., Pleij, C. W. A., Heus, H. A., and Hilbers, C. W. (1998) *Science* **280**, 434–438

327. Shi, P.-Y., Brinton, M. A., Veal, J. M., Zhong, Y. Y., and Wilson, W. D. (1996) *Biochemistry* **35**, 4222–4230
328. Hubbard, J. M., and Hearst, J. E. (1991) *Biochemistry* **30**, 5458–5465
329. Soll, D., and Abelson, J., eds. (1979) *Transfer RNA*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
330. Zimmermann, R. A. (1996) *Science* **271**, 1240–1241
331. Steinberg, S., Gautheret, D., and Cedergren, R. (1994) *J. Mol. Biol.* **236**, 982–989
332. Holbrook, S. R., Sussman, J. L., Warrant, R. W., and Kim, S.-H. (1978) *J. Mol. Biol.* **123**, 631–660
333. Rich, A., and Kim, S. H. (1978) *Sci. Am.* **238** (Jan), 52–62
334. Rich, A. (1978) *Trends Biochem. Sci.* **3**, 34–37
335. Moras, D., Comarmond, M. B., Fischer, J., Weiss, R., Thierry, J. C., Ebel, J. P., and Giegé, R. (1980) *Nature (London)* **288**, 669–674
336. Woo, N. H., Roe, B. A., and Rich, A. (1980) *Nature (London)* **286**, 346–351
337. Wakao, H., Romby, P., Westhof, E., Laalami, S., Grunberg-Manago, M., Ebel, J.-P., Ehresmann, C., and Ehresmann, B. (1989) *J. Biol. Chem.* **264**, 20363–20371
338. Hou, Y.-M., and Schimmel, P. (1992) *Biochemistry* **31**, 4157–4160
339. Leehey, M. A., Squassoni, C. A., Friederich, M. W., Mills, J. B., and Hagerman, P. J. (1995) *Biochemistry* **34**, 16235–16239
340. Quigley, G. J., and Rich, A. (1976) *Science* **194**, 796–806
341. Brimacombe, R., Atmadja, J., Stiege, W., and Schüler, D. (1988) *J. Mol. Biol.* **199**, 115–136
342. Mueller, F., and Brimacombe, R. (1997) *J. Mol. Biol.* **271**, 524–544
343. Turner, D. H., and Sugimoto, N. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 167–192
344. Brimacombe, R. (1988) *Biochemistry* **27**, 4207–4213
345. Schnare, M. N., Damberger, S. H., Gray, M. W., and Gutell, R. R. (1996) *J. Mol. Biol.* **256**, 701–719
346. Yonath, A. (1984) *Trends Biochem. Sci.* **9**, 227–230
347. Pennisi, E. (1999) *Science* **285**, 2048–2051
348. Svergun, D. I., Koch, M. H. J., Pedersen, J. S., and Serdyuk, I. N. (1994) *J. Mol. Biol.* **240**, 78–86
349. Vladimirov, S. N., Druzina, Z., Wang, R., and Cooperman, B. S. (2000) *Biochemistry* **39**, 183–193
350. Gregory, S. T., and Dahlberg, A. E. (1999) *J. Mol. Biol.* **285**, 1475–1483
351. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) *Science* **289**, 905–920
352. Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N., and Noller, H. F. (1999) *Science* **285**, 2095–2104
353. Culver, G. M., Cate, J. H., Yusupova, G. Z., Yusupov, M. M., and Noller, H. F. (1999) *Science* **285**, 2133–2135
354. Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000) *Nature (London)* **407**, 340–348
355. Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000) *Science* **289**, 920–930
356. Muth, G. W., Ortoleva-Donnelly, L., and Strobel, S. A. (2000) *Science* **289**, 947–950
357. Cech, T. R. (2000) *Science* **289**, 878–879
358. Pennisi, E. (1996) *Science* **273**, 574–575
359. O'Gara, M., Klimasauskas, S., Roberts, R. J., and Cheng, X. (1996) *J. Mol. Biol.* **261**, 634–645
360. Gommers-Ampt, J. H., and Borst, P. (1995) *FASEB J.* **9**, 1034–1042
361. Maltman, K. L., Neuhard, J., and Warren, R. A. J. (1981) *Biochemistry* **20**, 3586–3591
362. Lipscomb, L. A., Peek, M. E., Morningstar, M. L., Verghis, S. M., Miller, E. M., Rich, A., Essigmann, J. M., and Williams, L. D. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 719–723
363. Leonard, G. A., McAuley-Hecht, K. E., Gibson, N. J., Brown, T., Watson, W. P., and Hunter, W. N. (1994) *Biochemistry* **33**, 4755–4761
364. Wrzesinski, J., Bakin, A., Nurse, K., Lane, B. G., and Ofengand, J. (1995) *Biochemistry* **34**, 8904–8913
365. Ellington, A. D., and Szostak, J. W. (1990) *Nature (London)* **346**, 818–822
366. Fan, P., Suri, A. K., Fiala, R., Live, D., and Patel, D. J. (1996) *J. Mol. Biol.* **258**, 480–500
367. Jenison, R. D., Gill, S. C., Pardi, A., and Polisky, B. (1994) *Science* **263**, 1425–1429
368. Hamasaki, K., Killian, J., Cho, J., and Rando, R. R. (1998) *Biochemistry* **37**, 656–663
369. Connell, G. J., Illangesekare, M., and Yarus, M. (1993) *Biochemistry* **32**, 5497–5502
370. Patel, D. J., Suri, A. K., Jiang, F., Jiang, L., Fan, P., Kumar, R. A., and Nonin, S. (1997) *J. Mol. Biol.* **272**, 645–664
371. Crick, F. H. C. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**, 3–9
372. Cornish-Bowden, A. (1996) *Trends Biochem. Sci.* **21**, 155
373. Hall, B. D. (1979) *Nature (London)* **282**, 129–130
374. Barrell, B. G., Bankier, A. T., and Drouin, J. (1979) *Nature (London)* **282**, 189–194
375. Horowitz, S., and Gorovsky, M. A. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 2452–2455
376. Chambers, I., and Harrison, P. R. (1987) *Trends Biochem. Sci.* **12**, 255–256
377. Gardner, M. (1970) *Sci. Am.* **223**(Aug), 110–112
378. Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., and Frank-Kamenetskii, M. D. (1987) *Nature (London)* **330**, 495–497
379. Guieysse, A.-L., Praseuth, D., and Hélène, C. (1997) *J. Mol. Biol.* **267**, 289–298
380. Panayotatos, N., and Fontaine, A. (1987) *J. Biol. Chem.* **262**, 11364–11368
381. Pabo, C. O., and Sauer, R. T. (1984) *Ann. Rev. Biochem.* **53**, 293–321
382. Berg, O. G., and von Hippel, P. H. (1988) *Trends Biochem. Sci.* **13**, 207–211
383. Marmorstein, R. Q., Joachimiak, A., Sprinzl, M., and Sigler, P. B. (1987) *J. Biol. Chem.* **262**, 4922–4927
384. Muto, A., and Osawa, S. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 166–169
385. Singer, C. E., and Ames, B. N. (1979) *Science* **170**, 822–826
386. Singer, C. E., and Ames, B. N. (1972) *Science* **175**, 1391
387. Harrison, S. C. (1991) *Nature (London)* **353**, 715–719
388. Churchill, M. E. A., and Travers, A. A. (1991) *Trends Biochem. Sci.* **16**, 92–97
389. Travers, A. (1993) *DNA-Protein Interactions*, Chapman & Hall, New York
390. Jones, S., van Heyningen, P., Berman, H. M., and Thornton, J. M. (1999) *J. Mol. Biol.* **287**, 877–896
391. Arents, G., and Moudrianakis, E. N. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11170–11174
392. Starich, M. R., Sandman, K., Reeve, J. N., and Summers, M. F. (1996) *J. Mol. Biol.* **255**, 187–203
393. Hud, N. V., Milanovich, F. P., and Balhorn, R. (1994) *Biochemistry* **33**, 7528–7535
394. Lohman, T. M., and Ferrari, M. E. (1994) *Ann. Rev. Biochem.* **63**, 527–570
395. Overman, L. B., and Lohman, T. M. (1994) *J. Mol. Biol.* **236**, 165–178
396. Folmer, R. H. A., Nilges, M., Folkers, P. J. M., Konings, R. N. H., and Hilbers, C. W. (1994) *J. Mol. Biol.* **240**, 341–357
397. Olah, G. A., Gray, D. M., Gray, C. W., Kergil, D. L., Sosnick, T. R., Mark, B. L., Vaughan, M. R., and Trewhella, J. (1995) *J. Mol. Biol.* **249**, 576–594
398. Kanellopoulos, P. N., Tsernoglou, D., van der Vliet, P. C., and Tucker, P. A. (1996) *J. Mol. Biol.* **257**, 1–8
399. Bogdarina, I., Fox, D. G., and Kneale, G. G. (1998) *J. Mol. Biol.* **275**, 443–452
400. Pluta, A. F., Mackay, A. M., Ainsztein, A. M., Goldberg, I. G., and Earnshaw, W. C. (1995) *Science* **270**, 1591–1594
401. Schevitz, R. W., Otwinowski, Z., Joachimiak, A., Lawson, C. L., and Sigler, P. B. (1985) *Nature (London)* **317**, 782–786
402. Zhang, R.-G, Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., and Sigler, P. B. (1987) *Nature (London)* **327**, 591–597

## References


403. Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., and Sigler, P. B. (1988) *Nature (London)* **335**, 321–329
404. Shakked, Z., Guzikevich– Guerstein, G., Frolow, F., Rabinovich, D., Joachimiak, A., and Sigler, P. B. (1994) *Nature (London)* **368**, 469–473
405. Lawson, C. L., and Carey, J. (1993) *Nature (London)* **366**, 178–182
406. Lawson, C. L., and Sigler, P. B. (1988) *Nature (London)* **333**, 869–871
407. Guenot, J., Fletterick, R. J., and Kollman, P. A. (1994) *Protein Sci.* **3**, 1276–1285
408. Rastinejad, F., Artz, P., and Lu, P. (1993) *J. Mol. Biol.* **233**, 389–399
409. Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S., and Miller, J. H. (1994) *J. Mol. Biol.* **240**, 421–433
410. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996) *Science* **271**, 1247–1254
411. Brennan, R. G., and Matthews, B. W. (1989) *J. Biol. Chem.* **264**, 1903–1906
412. Jordan, S. R., and Pabo, C. O. (1988) *Science* **242**, 893–899
413. Wolberger, C., Dong, Y., Ptashne, M., and Harrison, S. C. (1988) *Nature (London)* **335**, 789–795
414. Aggarwal, A. K., Rodgers, D. W., Drottar, M., Ptashne, M., and Harrison, S. C. (1988) *Science* **242**, 899–907
415. Bell, A. C., and Koudelka, G. B. (1993) *J. Mol. Biol.* **234**, 542–553
416. Mandel-Gutfreund, Y., Schueler, O., and Margalit, H. (1995) *J. Mol. Biol.* **253**, 370–382
417. Ferré-D'Amaré, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993) *Nature (London)* **363**, 38–44
418. Junius, F. K., O'Donoghue, S. I., Nilges, M., Weiss, A. S., and King, G. F. (1996) *J. Biol. Chem.* **271**, 13663–13667
419. Keller, W., König, P., and Richmond, T. J. (1995) *J. Mol. Biol.* **254**, 657–667
420. Harbury, P. B., Kim, P. S., and Alber, T. (1994) *Nature (London)* **371**, 80–83
421. Klug, A., and Schwabe, J. W. R. (1995) *FASEB J.* **9**, 597–604
422. Rhodes, D., and Klug, A. (1993) *Sci. Am.* **268** (Feb), 56–65
423. Pavletich, N. P., and Pabo, C. O. (1991) *Science* **252**, 809–817
424. Mackay, J. P., and Crossley, M. (1998) *Trends Biochem. Sci.* **23**, 1–4
425. Miller, J., McLachlan, A. D., and Klug, A. (1985) *EMBO J.* **4**, 1609–1614
426. Pieler, T., and Theunissen, O. (1993) *Trends Biochem. Sci.* **18**, 226–230
427. Pavletich, N. P., and Pabo, C. O. (1993) *Science* **261**, 1701–1707
428. Krizek, B. A., Amann, B. T., Kilfoil, V. J., Merkle, D. L., and Berg, J. M. (1991) *J. Am. Chem. Soc.* **113**, 4518–4523
429. Somers, W. S., and Phillips, S. E. V. (1992) *Nature (London)* **359**, 387–393
430. Phillips, S. E. V. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 671–701
431. Breg, J. N., van Opheusden, H. J., Burgering, M. J. M., Boelens, R., and Kaptein, R. (1990) *Nature (London)* **346**, 586–589
432. Bonnefoy, E., Takahashi, M., and Yaniv, J. R. (1994) *J. Mol. Biol.* **242**, 116–129
433. Müller, C. W., Rey, F. A., Sodeoka, M., Verdine, G. L., and Harrison, S. C. (1995) *Nature (London)* **373**, 311–317
434. Ghosh, G., Van Duyne, G., Ghosh, S., and Sigler, P. B. (1995) *Nature (London)* **373**, 303–310
435. Baltimore, D., and Beg, A. A. (1995) *Nature (London)* **373**, 287–288
436. Kaufmann, E., Müller, D., and Knöchel, W. (1995) *J. Mol. Biol.* **248**, 239–254
437. Burd, C. G., and Dreyfuss, G. (1994) *Science* **265**, 615–621
438. Nagai, K., and Mattaj, I. W., eds. (1994) *RNA-Protein Interactions*, IRL Press, Oxford
439. Draper, D. E. (1995) *Ann. Rev. Biochem.* **64**, 593–620
440. Hentze, M. W. (1994) *Trends Biochem. Sci.* **19**, 101–103
441. Allain, F. H.-T., Gubser, C. C., Howe, P. W. A., Nagai, K., Neuhaus, D., and Varani, G. (1996) *Nature (London)* **380**, 646–650
442. Kenan, D. J., Query, C. C., and Keene, J. D. (1991) *Trends Biochem. Sci.* **16**, 214–220
443. Fierro-Monti, I., and Mathews, M. B. (2000) *Trends Biochem. Sci.* **25**, 241–246
444. Voyles, B. A. (1993) *The Biology of Viruses*, Mosby, St. Louis, Missouri
445. Fields, B. N., and Knipe, D. M., eds. (1990) *Fields Virology*, 2nd ed., Raven Press, New York
446. Cann, A. (1993) *Principles of Molecular Virology*, Academic Press, San Diego, California
447. Endemann, H., and Model, P. (1995) *J. Mol. Biol.* **250**, 496–506
448. Williams, K. A., Glibowicka, M., Li, Z., Li, H., Khan, A. R., Chen, Y. M. Y., Wang, J., Marvin, D. A., and Deber, C. M. (1995) *J. Mol. Biol.* **252**, 6–14
449. Wen, Z. Q., Overman, S. A., and Thomas, G. J., Jr. (1997) *Biochemistry* **36**, 7810–7820
450. Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., III, Slocombe, P. M., and Smith, M. (1977) *Nature (London)* **265**, 687–695
451. McKenna, R., Ilag, L. L., and Rossmann, M. G. (1994) *J. Mol. Biol.* **237**, 517–543
452. McKenna, R., Xia, D., Willingmann, P., Ilag, L. L., Krishnaswamy, S., Rossmann, M. G., Olson, N. H., Baker, T. S., and Incardona, N. L. (1992) *Nature (London)* **355**, 137–143
453. Godson, G. N. (1978) *Trends Biochem. Sci.* **3**, 249–253
454. McKenna, R., Bowman, B. R., Ilag, L. L., Rossmann, M. G., and Fane, B. A. (1996) *J. Mol. Biol.* **256**, 736–750
455. Wu, H., and Rossmann, M. G. (1993) *J. Mol. Biol.* **233**, 231–244
456. Tsao, J., Chapman, M. S., Agbandje, M., Keller, W., Smith, K., Wu, H., Luo, M., Smith, T. J., Rossmann, M. G., Compans, R. W., and Parrish, C. R. (1991) *Science* **251**, 1456–1464
457. Chipman, P. R., Agbandje-McKenna, M., Kajigaya, S., Brown, K. E., Young, N. S., Baker, T. S., and Rossmann, M. G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7502–7506
458. Howarth, A. J., and Goodman, R. M. (1982) *Trends Biochem. Sci.* **7**, 180–182
459. Tiollais, P., Pourcel, C., and Dejean, A. (1985) *Nature (London)* **317**, 489–495
460. Tiollais, P., and Buendia, M.-A. (1991) *Sci. Am.* **264**(Apr), 116–123
461. Böttcher, B., Wynne, S. A., and Crowther, R. A. (1997) *Nature (London)* **386**, 88–91
462. Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L., and Weissman, S. M. (1978) *Science* **200**, 494–502
463. Liddington, R. C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T. L., and Harrison, S. C. (1991) *Nature (London)* **354**, 278–284
464. Griffith, J. P., Griffith, D. L., Rayment, I., Murakami, W. T., and Caspar, D. L. D. (1992) *Nature (London)* **355**, 652–654
465. Yang, R. C. A., and Wu, R. (1979) *Science* **206**, 456–462
466. Howley, P. M. (1986) *N. Engl. J. Med.* **315**, 1089–1090
467. Hess, M., Cuzange, A., Ruigrok, R. W. H., Chroboczek, J., and Jacrot, B. (1995) *J. Mol. Biol.* **252**, 379–385
468. Athappilly, F. K., Murali, R., Rux, J. J., Cai, Z., and Burnett, R. M. (1994) *J. Mol. Biol.* **242**, 430–455
469. Witkowski, J. A. (1988) *Trends Biochem. Sci.* **13**, 110–113
470. Sugden, B. (1991) *Trends Biochem. Sci.* **16**, 45–46
471. Zhou, Z. H., Prasad, B. V. V., Jakana, J., Rixon, F. J., and Chiu, W. (1994) *J. Mol. Biol.* **242**, 456–469
472. Trus, B. L., Booy, F. P., Newcomb, W. W., Brown, J. C., Homa, F. L., Thomsen, D. R., and Steven, A. C. (1996) *J. Mol. Biol.* **263**, 447–462
473. Henle, W., Henle, G., and Lennette, E. T. (1974) *Sci. Am.* **241**(Jul), 48–59
474. Berardelli, P. (1997) *Science* **278**, 1710
475. Senkevich, T. G., Bugert, J. J., Sisler, J. R., Koonin, E. V., Darai, G., and Moss, B. (1996) *Science* **273**, 813–816
476. Hohn, T., Hohn, B., and Pfeiffer, P. (1985) *Trends Biochem. Sci.* **10**, 205–209
477. Rosa, P., Mantovani, S., Rosboch, R., and Huttner, W. B. (1992) *J. Biol. Chem.* **267**, 12227–12232
478. Diener, T. O. (1991) *FASEB J.* **5**, 2808–2813
479. Diener, T. V. (1984) *Trends Biochem. Sci.* **9**, 133–136
480. Diener, T. O. (1979) *Viroids and Viroid Diseases*, Wiley, New York
481. Martínez-Soriano, J. P., Galindo-Alonso, J., Maroon, C. J. M., Yucel, I., Smith, D. R., and Diener, T. O. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9397–9401
482. Polson, A. G., Bass, B. L., and Casey, J. L. (1996) *Nature (London)* **380**, 454–456
483. Wang, K.-S., Choo, Q.-L., Weiner, A. J., Ou, J.-H., Najarian, R. C., Thayer, R. M., Mullenbach, G. T., Denniston, K. J., Gerin, J. L., and Houghton, M. (1986) *Nature (London)* **323**, 508–514
484. Lai, M. M. C. (1995) *Ann. Rev. Biochem.* **64**, 259–286
485. Unge, T., Liljas, L., Strandberg, B., Vaara, I., Kannan, K. K., Fridborg, K., Nordman, C. E., and Lentz, P. J., Jr. (1980) *Nature (London)* **285**, 373–377
486. Larson, S. B., Koszelak, S., Day, J., Greenwood, A., Dodds, J. A., and McPherson, A. (1993) *Nature (London)* **361**, 179–182
487. Larson, S. B., Koszelak, S., Day, J., Greenwood, A., Dodds, J. A., and McPherson, A. (1993) *J. Mol. Biol.* **231**, 375–391
488. Golmohammadi, R., Valegård, K., Fridborg, K., and Liljas, L. (1993) *J. Mol. Biol.* **234**, 620–639
489. LeCuyer, K. A., Behlen, L. S., and Uhlenbeck, O. C. (1995) *Biochemistry* **34**, 10600–10606
490. Stonehouse, N. J., Valegård, K., Golmohammadi, R., van den Worm, S., Walton, C., Stockley, P. G., and Liljas, L. (1996) *J. Mol. Biol.* **256**, 330–339
491. Beremarand, M. N., and Blumenthal, T. (1979) *Cell* **18**, 257–266
492. Hogle, J. M., Chow, M., and Filman, D. J. (1985) *Science* **229**, 1358–1365
493. Hogle, J. M., Chow, M., and Filman, D. J. (1987) *Sci. Am.* **256**(Mar), 42–49
494. Najaran, R., Caput, D., Gee, W., Potter, S. J., Renard, A., Merryweather, J., Van Nest, G., and Dina, D. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 2627–2631
495. Rossmann, M. G., Arnold, E., Griffith, J. P., Kamer, G., Luo, M., Smith, T. J., Vriend, G., Rueckert, R. R., Sherry, B., McKinlay, M. A., Diana, G., and Otto, M. (1987) *Trends Biochem. Sci.* **12**, 313–318

## References


496. Olson, N. H., Kolatkar, P. R., Oliveira, M. A., Cheng, R. H., Greve, J. M., McClelland, A., Baker, T. S., and Rossmann, M. G. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 507–511
497. Hadfield, A. T., Oliveira, M. A., Kim, K. H., Minor, I., Kremer, M. J., Heinz, B. A., Shepard, D., Pevear, D. C., Rueckert, R. R., and Rossmann, M. G. (1995) *J. Mol. Biol.* **253**, 61–73
498. Küpper, H., Keller, W., Kurz, C., Forss, S., Schaller, H., Franze, R., Stohmaier, K., Marquardt, O., Zaslavsky, V. G., and Hofschneider, P. H. (1981) *Nature (London)* **289**, 555–559
499. Brown, F. (1981) *Trends Biochem. Sci.* **6**, 325–327
500. Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D., and Brown, F. (1989) *Nature (London)* **337**, 709–716
501. Luo, M., Vriend, G., Kamer, G., Minor, I., Arnold, E., Rossmann, M. G., Boege, U., Scraba, D. G., Duke, G. M., and Palmenberg, A. C. (1987) *Science* **235**, 182–191
502. Choi, H.-K., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M. G., and Wengler, G. (1991) *Nature (London)* **354**, 37–43
503. Simons, K., Garoff, H., and Helenius, A. (1982) *Sci. Am.* **246**(Feb), 58–66
504. Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K., and Bricogne, G. (1978) *Nature (London)* **276**, 368–373
505. Erickson, J. W., Silva, A. M., Murthy, M. R. N., Fita, I., and Rossmann, M. G. (1985) *Science* **229**, 625–629
506. Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R., and Klug, A. (1978) *Nature (London)* **276**, 362–368
507. Holmes, K. C. (1980) *Trends Biochem. Sci.* **5**, 4–7
507a. Lauffer, M. A. (1984) *Trends Biochem. Sci.* **9**, 369–371
508. Wang, H., and Stubbs, G. (1994) *J. Mol. Biol.* **239**, 371–384
509. Kaplan, M. M., and Webster, R. G. (1977) *Sci. Am.* **237**(Dec), 88–106
510. von Itzstein, M., and 17 other authors. (1993) *Nature (London)* **363**, 418–423
511. Clarke, D. K., Duarte, E. A., Elena, S. F., Moya, A., Domingo, E., and Holland, J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4821–4824
512. Varmus, H. (1988) *Science* **240**, 1427–1434
513. Darlix, J.-L., Lapadat-Tapolsky, M., de Rocquigny, H., and Roques, B. P. (1995) *J. Mol. Biol.* **254**, 523–537
514. Löwer, R., Löwer, J., and Kurth, R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5177–5184
515. Stephenson, J. R., ed. (1980) *Molecular Biology of RNA Tumor Viruses*, Academic Press, New York
516. Compans, R. W., and Bishop, D. H. L., eds. (1983) *Double-Stranded RNA Viruses*, Elsevier, Biomedical, New York
517. Thornton, A., and Zuckerman, A. J. (1975) *Nature (London)* **254**, 557–558
518. Baldwin, M. A., Cohen, F. E., and Prusiner, S. B. (1995) *J. Biol. Chem.* **270**, 19197–19200
519. Manuelidis, L., Sklaviadis, T., Akowitz, A., and Fritch, W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5124–5128
520. Prusiner, S. B. (1996) *Trends Biochem. Sci.* **21**, 482–487
521. Prusiner, S. B. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13363–13383
522. Weissmann, C. (1999) *J. Biol. Chem.* **274**, 3–6
523. Chesebro, B. (1998) *Science* **286**, 660–662
524. Balter, M. (1999) *Science* **286**, 660–662
525. Temin, H. M. (1978) *Trends Biochem. Sci.* **3**, N80–N83
525a. Sambrook, J., and Russell, D. (2000) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Lab. Press, Plainview, New York
526. Grossman, L., and Moldave, K., eds. (1967) *Methods in Enzymology*, Vol. 12A, Academic Press, New York (pp. 531–708)
527. Robyt, J. F., and White, B. J. (1996) *Biochemical Techniques, Theory and Practice*, 2nd ed., Waveland Press, Prospect Heights, Illinois
528. Blackburn, G. M., and Gait, M. J., eds. (1996) *Nucleic Acids in Chemistry and Biology*, 2nd ed., Oxford Univ. Press, Oxford
529. Rajeswari, M. R., Montenay-Garestier, T., and Helene, C. (1987) *Biochemistry* **26**, 6825–6831
530. Hecker, R., Colpan, M., and Riesner, D. (1985) *Journal of Chromatography* **326**, 251–261
531. Edwardson, P. A. D., Atkinson, T., Lowe, C. R., and Small, D. A. P. (1986) *Anal. Biochem.* **152**, 215–220
532. Lawrance, S. K., Smith, C. L., Srivastava, R., Cantor, C. R., and Weissman, S. M. (1987) *Science* **235**, 1387–1390
533. Cantor, C. R., Smith, C. L., and Matthew, M. K. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 287–304
534. Smith, C. L., and Cantor, C. R. (1987) *Trends Biochem. Sci.* **12**, 284–287
535. Kochetkov, N. K., and Budovskii, E. I. (1972) *Organic Chemistry of Nucleic Acids*, Plenum, New York (Part A, pp. 137–147)
536. Bozic, D., Grazulis, S., Siksnys, V., and Huber, R. (1996) *J. Mol. Biol.* **255**, 176–186
537. Smith, H. O. (1979) *Science* **205**, 455–462
538. Air, G. M. (1979) *CRC Crit. Revs. Biochem.* **6**, 1–33
539. Roberts, R. J. (1980) *Methods Enzymol.* **65**, 1–15
540. Koob, M., and Szybalski, W. (1990) *Science* **250**, 271–273
541. Veselkov, A. G., Demidov, V. V., Frank-Kamenetskii, M. D., and Nielsen, P. E. (1996) *Nature (London)* **379**, 214
542. Zaug, A. J., Been, M. D., and Cech, T. R. (1986) *Nature (London)* **324**, 429–433
543. Uhlenbeck, O. C. (1987) *Nature (London)* **328**, 596–600
544. Landgraf, R., Chen, C.-hB., and Sigman, D. S. (1994) *Biochemistry* **33**, 10607–10615
545. Adams, R. L. P., Burdon, R. H., Campbell, A. M., and Smellie, R. M. S. (1976) *Davidson's The Biochemistry of the Nucleic Acids*, 8th ed., Academic Press, New York (pp. 50–82)
546. Rayford, R., Anthony, D. D., Jr., O'Neill, R. E., Jr., and Merrick, W. C. (1985) *J. Biol. Chem.* **260**, 15708–15713
547. Bochner, B. R., and Ames, B. N. (1982) *J. Biol. Chem.* **257**, 9759–9769
548. Chen, H., and Shaw, B. R. (1994) *Biochemistry* **33**, 4121–4129
549. Hayatsu, H., and Shiragami, M. (1979) *Biochemistry* **18**, 632–647
550. Spiro, T. G., ed. (1980) *Nucleic Acid-Metal Ion Interactions*, Wiley, New York
551. Dische, Z. (1955) *Nucleic Acids* **1**, 755
552. Marmur, J., and Doty, P. (1962) *J. Mol. Biol.* **5**, 109–118
553. Korolev, N., Lyubartsev, A. P., and Nordenskiöld, L. (1998) *Biophys. J.* **75**, 3041–3056
554. Davidson, J. N. (1972) *The Biochemistry of Nucleic Acids*, 7th ed., Academic Press, New York (p. 148)
555. Wada, A., Yabuki, S., and Husimi, Y. (1980) *Crit. Revs. Biochem.* **9**, 87–144
556. Britten, R. J., and Kohne, D. E. (1968) *Science* **161**, 529–540
557. Wilson, D. A., and Thomas, C. A., Jr. (1974) *J. Mol. Biol.* **84**, 115–144
558. Bendich, A. J., and Bolton, E. T. (1968) in *Methods in Enzymology*, Vol. 12B (Grossman, L., and Moldave, K., eds), pp. 635–640, Academic Press, New York
559. Gillespie, D. (1968) in *Methods in Enzymology*, Vol. 12B (Grossman, L., and Moldave, K., eds), pp. 641–668, Academic Press, New York
560. Hall, B. D., Haarr, L., and Kleppe, K. (1980) *Trends Biochem. Sci.* **5**, 254–256
561. Britten, R. J., and Kohne, D. E. (1967) *Carnegie Instit. Washington Yearbook*, 65th ed., Carnegie Instit., Washington, D. C. (pp. 78–106)
562. Marmur, J. (1994) *Trends Biochem. Sci.* **19**, 343–346
563. Szabo, P., and Ward, D. C. (1982) *Trends Biochem. Sci.* **7**, 425–427
564. Kessler, C., ed. (1992) *Nonradioactive Labeling and Detection of Biomolecules*, Springer-Verlag, New York
565. Leitch, A. R., Chwarzacher, T., Jackson, D., and Leitch, I. J. (1994) *In Situ Hybridization: a Practical Guide*, Bios, Oxford
566. Kumar, A., Tchen, P., Roullet, F., and Cohen, J. (1988) *Anal. Biochem.* **169**, 376–
567. Isaac, P. G., ed. (1994) *Protocols for Nucleic Acid Analysis by Nonradioactive Probes*, Humana Press, Totowa, New Jersey
568. Gillevet, P. M. (1990) *Nature (London)* **348**, 657–658
569. Negro, F., Pacchioni, D., Shimizu, Y., Miller, R. H., Bussolati, G., Purcell, R. H., and Bonino, F. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2247–2251
570. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York
571. Gubler, U., and Hoffman, B. J. (1983) *Gene* **25**, 263–269
572. Watson, C. J., and Jackson, J. F. (1985) in *DNA Cloning Vol. I: A Practical Approach* (Glover, D. M., ed), pp. 79–100, IRL Press, Washington, DC
573. Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y., and Matsubara, K. (1985) *J. Biol. Chem.* **260**, 2605–2608
574. Warner, B. D., Warner, M. E., Karus, G. A., Ku, L., Brown-Shimer, S., and Urdea, M. S. (1983) *DNA* **4**, 401–411
575. Sekiya, T., Takeya, T., Brown, E. L., Belagaje, R., Contreras, R., Fritz, H.-J., Gait, M. J., Lees, R. G., Ryan, M. J., Khorana, H. G., and Norris, K. E. (1979) *J. Biol. Chem.* **254**, 5787–5801
576. Wada, T., Sato, Y., Honda, F., Kawahara, S.-i, and Sekine, M. (1997) *J. Am. Chem. Soc.* **119**, 12710–12721
577. Matteucci, M. D., and Caruthers, M. H. (1981) *J. Am. Chem. Soc.* **103**, 3185–3191
578. Beaucage, S. L., and Caruthers, M. H. (1981) *Tetrahedron Letters* **22**, 1859–1862
579. Air, G. M. (1979) *CRC Crit. Revs. Biochem.* **6**, 1–33
580. Alvarado-Urbina, G., Sathe, G. M., Liu, W.-C., Gillen, M. F., Duck, P. D., Bender, R., and Ogilvie, K. K. (1981) *Science* **214**, 270–274
581. Caruthers, M. H. (1985) *Science* **230**, 281–285
582. Lashkari, D. A., Hunicke-Smith, S. P., Norgren, R. M., Davis, R. W., and Brennan, T. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7912–7915
583. Usman, N., Ogilvie, K. K., Jiang, M.-Y., and Cedergren, R. J. (1987) *J. Am. Chem. Soc.* **109**, 7845–7854
584. Ebe, K., Schöed, M., Rossi, J. J., and Wallace, R. B. (1987) *DNA* **6**, 497–504
585. Usman, N., and Cedergren, R. (1992) *Trends Biochem. Sci.* **17**, 334–339
586. Eaton, B. E., and Pieken, W. A. (1995) *Ann. Rev. Biochem.* **64**, 837–863
587. Rohatgi, R., Bartel, D. P., and Szostak, J. W. (1996) *J. Am. Chem. Soc.* **118**, 3340–3344
588. Kleppe, K., Ohtsuka, E., Kleppe, R., and Khorana, H. G. (1971) *J. Mol. Biol.* **56**, 341–361
589. Panet, A., and Khorana, H. G. (1974) *J. Biol. Chem.* **249**, 5213–5221

## References


590. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985) *Science* **230**, 1350–1354
591. Mullis, K. B. (1990) *Sci. Am.* **262**(Apr), 56–65
592. Rabinow, P. (1996) *Making PCR: A Story of Biotechnology*, Univ. Chicago Press, Chicago, Illinois
593. Mullis, K., Ferré, F., and Gibbs, R., eds. (1994) *The Polymerase Chain Reaction*, Birkhäuser, Boston, Massachusetts
594. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) *Science* **239**, 487–491
595. Erlich, H. A., Gelfand, D. H., and Saiki, R. K. (1988) *Nature (London)* **331**, 461–462
596. Appenzeller, T. (1990) *Science* **247**, 1030–1032
597. Newton, C., and Graham, A. (1997) *PCR*, 2nd ed., Springer-Verlag, New York
598. Bloch, W. (1991) *Biochemistry* **30**, 2735–2747
599. White, B. A., ed. (1993) *PCR Protocols*, Humana Press, Totowa, New Jersey

599a. Innis, M. A., Gelfand, D. H., and Sninsky, J. J., eds. (1999) *PCR Applications*, Academic Press, San Diego, California

600. Cohen, J. (1994) *Science* **263**, 1564–1565
601. Pääbo, S. (1993) *Sci. Am.* **269**(Nov), 86–92
602. Pääbo, S., Higuchi, R. G., and Wilson, A. C. (1989) *J. Biol. Chem.* **264**, 9709–9712
603. Höss, M., Pääbo, S., and Vereshchagin, N. K. (1994) *Nature (London)* **370**, 333
604. Scharf, S. J., Horn, G. T., and Erlich, H. A. (1986) *Science* **233**, 1076–1078
605. Li, H., Gyllensten, U. B., Cui, X., Saiki, R. K., Erlich, H. A., and Arnheim, N. (1988) *Nature (London)* **335**, 414–417
606. Engelke, D. R., Hoener, P. A., and Collins, F. S. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 544–548
607. Karlovsky, P. (1990) *Trends Biochem. Sci.* **15**, 419
608. Hengen, P. N. (1995) *Trends Biochem. Sci.* **20**, 372–373
609. Wong, G. K.-S., Yu, J., Thayer, E. C., and Olson, M. V. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5225–5230
610. Bernardi, G. (1978) *Nature (London)* **276**, 558–559
611. Bonitz, S. G., Coruzzi, G., Thalenfeld, B. E., Tzagoloff, A., and Macino, G. (1980) *J. Biol. Chem.* **255**, 11927–11941
612. Martin, N. C., Miller, D. L., and Donelson, J. E. (1979) *J. Biol. Chem.* **254**, 11729–11734
613. Fischer, S. G., and Lerman, L. S. (1979) *Methods Enzymol.* **68**, 183–191
614. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517
615. Southern, E. (1979) *Methods Enzymol.* **68**, 152–176
616. Grunstein, M., and Wallis, J. (1979) *Methods Enzymol.* **68**, 379–389
617. Sanger, F., and Coulson, A. R. (1975) *J. Mol. Biol.* **94**, 441–448
618. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
619. Alphey, L. (1997) *DNA Sequencing*, Springer-Verlag, New York
620. Ansorge, W., Voss, H., and Zimmerman, J., eds. (1997) *DNA Sequencing Strategies - Automated and Advanced Approaches*, Wiley, New York
621. Griffin, H. G., and Griffin, A. M., eds. (1993) *DNA Sequencing Protocols*, Humana Press, Totowa, New Jersey
622. Smith, A. J. H. (1980) *Methods Enzymol.* **65**, 560–580
623. Gillam, S., and Smith, M. (1980) *Methods Enzymol.* **65**, 687–701
624. Chen, E. Y., and Seeburg, P. H. (1985) *DNA* **4**, 165–170
625. Ornstein, D., Moen, P. T., and Kashdan, M. A. (1985) *DNA* **4**, 94
626. Bassam, B. J., Caetano-Anolles, G., and Gresshoff, P. M. (1991) *Anal. Biochem.* **196**, 80–83
627. Mizusawa, S., Nishimura, S., and Seela, F. (1986) *Nucleic Acids Res.* **14**, 1319–1324
628. Gerken, T. A., Gupta, R., and Jentoft, N. (1992) *Biochemistry* **31**, 639–648
629. Hengen, P. N. (1996) *Trends Biochem. Sci.* **21**, 33–34
630. Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E., and Surrey, S. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4298–4302
631. Henikoff, S. (1984) *Gene* **28**, 351–359
632. Adams, M. D., Kerlavage, A. R., Kelley, J. M., Gocayne, J. D., Fields, C., Fraser, C. M., and Venter, J. C. (1994) *Nature (London)* **368**, 474–475
633. Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H., and Hood, L. E. (1986) *Nature (London)* **321**, 674–679
634. Ju, J., Ruan, C., Fuller, C. W., Glazer, A. N., and Mathies, R. A. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4347–4351
635. Fu, D.-J., Broude, N. E., Köster, H., Smith, C. L., and Cantor, C. R. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10162–10166
636. Bult, C. J., and 39 other authors. (1996) *Science* **273**, 1058–1073
637. Sutcliffe, J. G. (1995) *Trends Biochem. Sci.* **20**, 87–90
638. Rowen, L., Mahairas, G., and Hood, L. (1997) *Science* **278**, 605–607
639. Mathies, R. A., Huang, X. C., and Quesada, M. A. (1992) *Anal. Chem.* **64**, 2149–2154
640. Yeung, E. S., and Li, Q. (1998) in *High Performance Capillary Electrophoresis* (Khaldi, M. G., ed), Wiley, New York
641. Little, D. P., Aaserud, D. J., Valaskovic, G. A., and McLafferty, F. W. (1996) *J. Am. Chem. Soc.* **118**, 9352–9359
642. Maxam, A. M., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 560–564
643. Gilbert, W. (1981) *Science* **214**, 1305–1312
644. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560
645. Brownlee, G. G. (1972) in *Determination of Sequences in RNA* (Work, T. S., and Work, E., eds), North-Holland Publ., Amsterdam
646. Pilly, D., Niemeyer, A., Schmidt, M., and Bargetzi, J. P. (1978) *J. Biol. Chem.* **253**, 437–445
647. Bayev, A. A. (1995) *Comprehensive Biochemistry* **38**, 439–479
648. Campbell, P. N. (1995) *Trends Biochem. Sci.* **20**, 259–260
649. Myers, T. W., and Gelfand, D. H. (1991) *Biochemistry* **30**, 7661–7666
650. Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R., and Guilley, H. (1977) *Nature (London)* **269**, 833–836
651. Stanley, J., and Vassilenko, S. (1978) *Nature (London)* **274**, 87–89
652. Kitamura, N., and Wimmer, E. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3196–3200
653. Josse, J., Kaiser, A. D., and Kornberg, A. (1961) *J. Biol. Chem.* **236**, 864–875
654. Stormo, G. D. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 241–263
655. Goad, W. B. (1986) *Ann. Rev. Biophys. Biophys. Chem.* **15**, 79–95
656. Garner, M. M., and Revzin, A. (1986) *Trends Biochem. Sci.* **11**, 395–396
657. Berger, R., Duncan, M. R., and Berman, B. (1993) *BioTechniques* **15**, 650–652
658. Welsh, J., and Cantor, C. R. (1984) *Trends Biochem. Sci.* **9**, 505–508
659. Galas, D., and Schmitz, A. (1978) *Nucleic Acids Res.* **5**, 3157–3170
660. Tullius, T. D. (1987) *Trends Biochem. Sci.* **12**, 297–300
661. Tullius, T. D., and Dombroski, B. A. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5469–5473
662. Prigodich, R. V., and Martin, C. T. (1990) *Biochemistry* **29**, 8017–8019
663. Flaus, A., Luger, K., Tan, S., and Richmond, T. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1370–1375
664. Tornaletti, S., and Pfeifer, G. P. (1995) *J. Mol. Biol.* **249**, 714–728

664a. Chaulk, S. G., Pezacki, J. P., and MacMillan, A. M. (2000) *Biochemistry* **39**, 10448–10453

665. Jaeger, J. A., SantaLucia, J., Jr., and Tinoco, I., Jr. (1993) *Ann. Rev. Biochem.* **62**, 255–287
666. Hüttenhofer, A., and Noller, H. F. (1994) *EMBO J.* **13**, 3892–3901
667. Laughrea, M., and Tam, J. (1992) *Biochemistry* **31**, 12035–12041
668. Reid, B. R. (1981) *Ann. Rev. Biochem.* **50**, 969–996
669. Hare, D. R., Ribeiro, N. S., Wemmer, D. E., and Reid, B. R. (1985) *Biochemistry* **24**, 4300–4306
670. Ikuta, S., Chattopadhyaya, R., Ito, H., Dickerson, R. E., and Kearns, D. R. (1986) *Biochemistry* **25**, 4840–4849
671. Lefèvre, J.-F., Lane, A. N., and Jardetzky, O. (1987) *Biochemistry* **26**, 5076–5090
672. Bax, A., and Lerner, L. (1986) *Science* **232**, 960–967
673. Patel, D. J., Shapiro, L., and Hare, D. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 423– 454
674. Flynn, P. F., Kintanar, A., Reid, B. R., and Drobny, G. (1988) *Biochemistry* **27**, 1191–1197
675. Cohen, J. S. (1987) *Trends Biochem. Sci.* **12**, 133–135
676. Wemmer, D. E., Chou, S.-H., and Reid, B. R. (1984) *J. Mol. Biol.* **180**, 41–60
677. Schmidt, P. G., Sierzputowska-Gracz, H., and Agris, P. F. (1987) *Biochemistry* **26**, 8529–8534
678. Farmer, B. T., Jr., and Müller, L. (1993) *J. Am. Chem. Soc.* **115**, 11040–11041
679. Nikonowicz, E. P., and Pardi, A. (1992) *Nature (London)* **355**, 184–186
680. Tate, S.-i, Ono, A., and Kainosho, M. (1994) *J. Am. Chem. Soc.* **116**, 5977–5978
681. Marino, J. P., Prestegard, J. H., and Crothers, D. M. (1994) *J. Am. Chem. Soc.* **116**, 2205–2206
682. Moore, P. B. (1995) *Acc. Chem. Res.* **28**, 251–256
683. Pardi, A. (1995) *Nature Struct. Biol.* **1**, 846–849
684. Mer, G., and Chazin, W. J. (1998) *J. Am. Chem. Soc.* **120**, 607–608

684a. Nikonowicz, E. P., Michnicka, M., and DeJong, E. (1998) *J. Am. Chem. Soc.* **120**, 3813–3814

684b. Liu, A., Majumdar, A., Hu, W., Kettani, A., Skripkin, E., and Patel, D. J. (2000) *J. Am. Chem. Soc.* **122**, 3206–3210

685. Tolbert, T. J., and Williamson, J. R. (1996) *J. Am. Chem. Soc.* **118**, 7929–7940
686. Xu, J., Lapham, J., and Crothers, D. M. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 44–48
687. Horowitz, J., Ofengand, J., Daniel, W. E. Jr., and Cohn, M. (1977) *J. Biol. Chem.* **252**, 4418–4420
688. Schroeder, S. A., Roongta, V., Fu, J. M., Jones, C. R., and Gorenstein, D. G. (1989) *Biochemistry* **28**, 8292–8303
689. Radhakrishnan, I., Patel, D. J., and Gao, X. (1992) *Biochemistry* **31**, 2514–2523
690. Schurter, E. J., Sayer, J. M., Oh-hara, T., Yeh, H. J. C., Yagi, H., Luxon, B. A., Jerina, D. M., and Gorenstein, D. G. (1995) *Biochemistry* **34**, 9009–9020
691. Mao, H., Deng, Z., Wang, F., Harris, T. M., and Stone, M. P. (1998) *Biochemistry* **37**, 4374–4387
692. Chou, S.-H., Zhu, L., Gao, Z., Cheng, J.-W., and Reid, B. R. (1996) *J. Mol. Biol.* **264**, 981–1001
693. Kettani, A., Guéron, M., and Leroy, J.-L. (1997) *J. Am. Chem. Soc.* **119**, 1108–1115
694. Chu, Wen-C., Liu, J. C.-H., and Horowitz, J. (1997) *Nucleic Acids Res.* **25**, 3944–3949

## Study Questions

1. Describe the typical distribution pattern of RNA and DNA in bacterial cells and in eukaryotic cells.

2. Draw the structures of the Watson–Crick base pairs guanine–cytosine (GC) and adenine–thymine (AT). Also draw the GU pair, which is not a Watson–Crick pair.

3. Draw the tautomeric structures possible for the cation formed by protonation of 9-methyladenine.

4. What unusual base pairs could arise from a minor tautomer of cytosine or from a minor tautomer of guanine?

5. The minor imino tautomeric form of adenosine occurs infrequently in DNA. Can this cause mutations? Explain; draw structures to illustrate your answer.

6. Will the substitution of hypoxanthine for adenine in DNA result in mutation? Explain.

7. Why is the methylation of DNA to form $O^6$-methyl guanine mutagenic?

8. Draw the structure of a dinucleotide that might be obtained by the partial hydrolysis of RNA. Indicate the following:
   a) The 5' end
   b) The 3' end
   c) The torsion angle $\chi$
   d) The point of cleavage by pancreatic ribonuclease
   e) The point of cleavage by periodic acid
   f) Two points at which the structure might be methylated by modifying enzymes acting on a polynucleotide

9. Draw the structure of guanosine-5'-phosphate in such a way that the configurations of the sugar ring and of the glycosidic linkage are clearly indicated. State whether you have drawn a *syn* or an *anti* conformer. Circle the most acidic proton in the guanine ring and indicate its approximate $pK_a$. Which is the most basic center? What is the approximate $pK_a$ of the conjugate acid?

10. What are the chemical functional groups in DNA? In RNA?

11. Electrophoresis of a mixture of the dinucleotides ApC and ApU at pH 3.5 separates two components. Identify these and explain the order of migration. Be as quantitative as possible.

12. Draw the structure of the predominant form of pGpC as it occurs at pH 3.5.

13. Why is DNA denatured at pH 11?

14. Draw a schematic representation of the polynucleotide portion of a DNA molecule and of an RNA molecule and indicate positions of cleavage by the following treatments:
   a) Mild HCl
   b) More vigorous HCl
   c) Mild NaOH
   d) More vigorous NaOH
   e) Pancreatic RNase
   f) Pancreatic DNase
   g) Splenic DNase
   h) Splenic phosphodiesterase
   i) Snake venom phophodiesterase
   j) Dnase from *Micrococcus*

15. A sample of DNA from a virus was hydrolyzed by acid and was found to have the following base composition (in mol%): adenine, 30; thymine, 39; guanine, 18; cytosine, 13. This differs from that of most DNA preparations. Offer a possible explanation. Sketch the expected temperature-absorbance profile of this DNA. Do you expect much hyperchromicity? Explain your answer.

16. Adenine is found to constitute 16.3% of the nucleic acid bases in a sample of bacterial DNA. What are the percentages of the other three bases?

17. For the following DNA sequence

    3'-CGATACGGCTATGCCATAGGC-5'

    write
    a) the sequence of the complementary DNA strand;
    b) the sequence of the corresponding segment of mRNA formed using the DNA segment above as the template;
    c) the amino acid sequence encoded by this segment.

18. What is the molecular mass of a segment of B-DNA that encodes a 386-residue protein? What is the length in nm? in Å? Do not make allowance for introns.

## Study Questions

19. Complete the following table:

| Name | Monomer | Linkage | Range of molecular masses |
|---|---|---|---|
| Protein | | | |
| Polysaccharide | | | |
| Nucleic acid | | | |
| Teichoic acid | | | |
| Poly-β-hydroxybutyrate | | | |

20. What is meant by the $T_m$ of a DNA sample? How does $T_m$ vary with base composition and what is the explanation of this?

21. Isolated "naked" bacterial DNA, from which proteins have been removed, is supercoiled. DNA in the bacterial chromosome is also supercoiled. When naked DNA is nicked, its supercoiling is abolished. In contrast nicking the chromosomal DNA does not abolish its supercoiling. Explain.

22. A closed circular duplex DNA has a 90 base-pair segment of alternating G and C residues. Upon transfer to a solution containing a high salt concentration, this segment undergoes a transition from the B conformation to the Z conformation.
    a) Explain why the high salt concentration induces a B→Z transformation.
    b) What changes would you expect in (1) the linking number *Lk*, (2) the writhe *Wr*, and (3) the twist *Tw* of the DNA as a result of this transition.

23. Name two or more characteristics of a DNA sequence or of its environment that will favor conversion of B-DNA into Z-DNA.

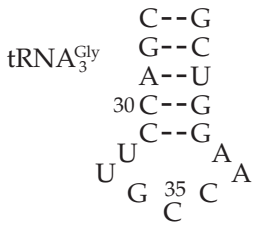
24. Suppose one double helical turn of a superhelical DNA molecule changes from a B conformation to the Z conformation. Calculate the approximate changes in (1) the linking $\Delta Lk$, (2) the writhe $\Delta Wr$, and (3) the twist $\Delta Tw$ of the DNA as a result of this transition. Show your calculations and explain your answers. For this problem assume that the B form of DNA has 10.4 bp per turn. Why is the B→Z transition favored in naturally occurring supercoiled DNA?

25. A circular DNA plasmid of length 1144 bp is supercoiled with a twist (*Tw*) of 110. Assume that the DNA has 10.4 bp per turn in its relaxed state.
    a) What is the linking number *Lk* and the writhe *Wr* in the plasmid?
    b) Is the plasmid negatively or positively supercoiled?
    c) Ethidium bromide is an intercalating agent that inserts between the stacked base pairs, separating the stacks and causing local unwinding that decreases the value of *Tw*. What effect would ethidium bromide have on the migration rate of the plasmid during electrophoresis?
    d) If part of the plasmid were to undergo a transition from B-DNA to Z-DNA, what would be the effect on *Lk*, *Tw*, and *Wr*?

26. You have been given a sample of nucleic acid, describe two ways you could determine whether it is RNA or DNA and two ways to determine whether it is single- or double-stranded.

27. What conclusions can you draw about the nature of the protein binding site on DNA from the observation that methylation of cytosine residues in the protein-recognition sequence inhibits protein binding?

28. The anticodon loop of one of the tRNA Gly molecules from *E. coli* is as follows. Identify the anticodon, reading from 3' to 5'. This tRNA recognizes two different Gly codons. What are they? Write them from 5' to 3'.




    The complete tRNA contains 75 nucleotides. Sketch the rest of the molecule in the cloverleaf representation. Label the 5' and 3' ends and the dihydrouridine and TψC loops. What are the last three nucleotides at the 3' end?

29. Are viruses alive? Explain your answer.


Ice and water are in equilibrium at 0°C and atmospheric pressure. When ice melts under these conditions the heat $Q$ absorbed from the surroundings is the **enthalpy change**, $\Delta H$, which equals 6.008 kJ mol$^{-1}$. For a reversible reaction the **entropy change**, $\Delta S$, equals $\Delta H/T$ and the entropy increases when the ice melts by $6.008 \times 10^3$ J / 273.16 = 22.0 JK$^{-1}$. Along the frozen edges of the river the ice and water are not in a true equilibrium but in a steady state. It is an **open system** in which water flows and energy is exchanged with the surroundings. Similarly, the cells of *Vorticella* represent open systems through which water and nutrients flow, and energy is exchanged with the surrounding water. Many chemical reactions within the cells are near equilibrium while the organism maintains its own structure. Photos: Frosty River, Banff Natl. Park, Alberta © Stephen J. Kraseman; *Vorticella* courtesy of Ralph Buchsbaum.

# Contents


281 **A. Thermodynamics**
282 1. The First Law of Thermodynamics
282 2. Enthalpy Changes and Thermochemistry
282 *Processes at constant pressure*
282 *Enthalpies of combustion and physiological fuel values*
284 3. The Second Law of Thermodynamics
284 *The thermodynamic temperature*
285 *Entropies from measurement of heat capacities*
285 4. A Criterion of Spontaneity: The Gibbs Energy
286 5. Practical Thermochemistry
286 *Summing changes in Gibbs energy*
287 *Reactions in solution*
287 *$\Delta G°$ and the equilibrium constant*
288 *Activity coefficients and concentration equilibrium constants*
289 *Changes in equilibria with temperature*
289 6. Thermodynamics and Life Processes
292 **B. Tables of $\Delta G°$ Values for Biochemical Compounds**
292 1. Gibbs Energies of Formation
292 2. Gibbs Energies of Dissociation of Protons
292 3. Group Transfer Potentials
293 4. "Constants" That Vary with pH and Magnesium Ion Concentrations
297 5. A New Standard for Biochemical Thermodynamic?
297 6. Bond Energies and Approximate Methods for Estimation of Thermodynamic Data
297 7. Gibbs Energies of Combustion by $O_2$ and by $NAD^+$
300 **C. Electrode Potentials and Gibbs Energy Changes for Oxidation–Reduction Reactions**
302 **D. The Adenylate System**
302 1. Storage and Utilization of Energy
303 2. Synthesis of ATP
303 3. Creatine Phosphate, an Energy Buffer
303 4. Phosphorus-31 NMR
304 **E. Complex Biochemical Equilibria**
305 1. Effects of pH on Equilbria
305 2. Microscopic Dissociation Constants and Tautomerization
307 3. The Binding of Metal Ions
307 *Factors affecting the strength of binding of a metal in a complex*
311 *How properties of the metal ion affect chelation*
311 *Metal binding sites in cells*
312 *Calcium-binding proteins*

319 **References**
321 **Study Questions**

**Boxes**

295 Box 6-A Measurement of Intracellular pH
299 Box 6-B Magnesium
308 Box 6-C Linear Gibbs Energy Relationships
314 Box 6-D Calcium
317 Box 6-E Metallothioneins

**Tables**

283 Table 6-1 Units of Energy and Work and the Values of Some Physical Constants
283 Table 6-2 Caloric Values of Food Components
285 Table 6-3 Entropies of Selected Substances
290 Table 6-4 Gibbs Energies of Formation and of Oxidation at 25°C for Compounds of Biochemical Interest
293 Table 6-5 Values of p$Ka$, $\Delta G°$, and $\Delta H°$ for Ionization of Acids at 25°C
294 Table 6-6 Gibbs Energies of Hydrolysis at 25°C (in kJ mol$^{-1}$)
298 Table 6-7 Empirical Bond Energies and Resonance Energies
301 Table 6-8 Reduction Potentials of Some Biologically Important Systems
310 Table 6-9 Logarithms of Binding Constants for Some 1:1 Metal Complexes at 25°C
311 Table 6-10 Ionic Radii in Nanometers for Some Metallic and Nonmetallic Ions

# Thermodynamics and Biochemical Equilibria

# 6

We all know from experience the importance of energy to life. We know that we must eat and that hard work not only tires us but also makes us hungry. Our bodies generate heat, an observation that led Lavoisier around 1780 to the conclusion that respiration represented a slow combustion of foods within the body. It soon became clear that respiration must provide the energy for both the mechanical work done by muscles and the chemical synthesis of body constituents. All organisms require energy and the ways in which living things obtain and utilize energy is a major theme of biology.

The discovery of the first and second laws of thermodynamics permitted the development of precise, quantitative relationships between heat, energy, and work. It also allowed **chemical equilibria** to be understood. Modern biochemical literature abounds with references to the thermodynamic quantities **energy** *E*, **enthalpy** *H*, **entropy** *S*, and **Gibbs energy** (also called free energy) *G*. The purposes of this chapter are: (1) to provide a short review of thermodynamic equations, (2) to provide tables of thermodynamic quantities for biochemical substances and to explain the use of these data in the consideration of equilibria in biochemical systems, and (3) to introduce the **adenylate system**, which consists of adenosine triphosphate (ATP), adenosine diphosphate (ADP), adenosine mono-phosphate (AMP), and inorganic phosphate ($P_i$). This system plays a central role in energy metabolism, and (4) to provide a quantitative understanding of the effects of pH and of metal ions on biochemical equilibria. *Many readers will want to go directly to Section D, which deals with the adenylate system and its significance for life.*

## A. Thermodynamics

Thermodynamics is concerned with the quantitative description of heat and energy changes and of chemical equilibria.[1–10] Knowledge of *changes* in thermodynamic quantities, such as $\Delta H$ and $\Delta S$, enables us to predict the equilibrium positions in reactions and whether or not under given circumstances a reaction will or will not take place. Furthermore, the consideration of thermodynamic quantities provides insight into the nature of forces responsible for bonding between molecules, enzymatic catalysis, functioning of DNA and RNA, and many other phenomena.

It is important to realize that while thermodynamic information will tell us whether or not a reaction can take place it says nothing about the rate of the reaction. It will not even say whether a reaction will proceed at all within a given period of time. This has led to the occasional assertion that thermodynamics is not relevant to biochemistry. This is certainly not true; it is important to understand energy relationships in biochemical reactions. At the same time, one should avoid the trap of assuming that thermodynamic calculations appropriate for *equilibrium* situations can always be applied directly to the *steady state* found in a living cell.

Thermodynamics is an exact science and its laws deal with measurable quantities whose values are determined only by the **state** of the **system** under consideration. For example, the system might be the solution in a flask resting in a thermostated bath. To specify its state we would have to say whether it is pure solid, liquid, or gas, or a solution of specified composition and give the temperature and pressure. The flask, the bath, and everything else would be

called **surroundings** or **environment**. The system plus surroundings is sometimes referred to as **the universe**.

## 1. The First Law of Thermodynamics

The first law of thermodynamics asserts the conservation of energy and also the equivalence of **work** and **heat**. Work and heat are both regarded as energy in transit. Heat may be absorbed by a system from the surroundings or evolved by a system and absorbed in the surroundings. Work can be done *by* a system on the surroundings or it can be done *on* a system. The first law postulates that there is an **internal energy** $E$ (also designated $U$), which is dependent only on the present state of the system and in no way is dependent upon the history of the system. The first law states that *E can be changed only by the flow of energy as heat or by work*. In other words, energy can neither be created nor destroyed.

In mathematical form, the first law is given as follows:

$$\Delta E = E\,(\text{products}) - E\,(\text{reactants}) = Q - W \quad (6\text{-}1)$$

Here $Q$ is the heat absorbed by the system from the surroundings and $W$ is the work done by the system on the surroundings. Energy, heat, and work are all measured in the same units. Chemists have traditionally used the **calorie** (cal) or **kilocalorie** (kcal) but are switching to the SI unit, the **joule** (Table 6-1).[10a] Work done by the system may be **mechanical** (e.g., by changing the volume of the surroundings), **electrical** (e.g., by charging of a battery), or **chemical** (e.g., by effecting the synthesis of a polypeptide from amino acids).

## 2. Enthalpy Changes and Thermochemistry

We are most often interested in the *changes* in the thermodynamic functions when a chemical reaction takes place; for example, the heat absorbed by the system within a **bomb calorimeter** where the volume stays constant ($Q_v$) is a direct measure of the change in $E$:

$$Q_V = \Delta E \quad (6\text{-}2)$$

To measure $\Delta E$ for combustion of a biochemical compound, the substance may be placed in a bomb together with gaseous oxygen and the mixture ignited within the calorimeter by an electric spark. In this case, heat will be evolved from the bomb and will pass into the surroundings. $Q_v$ and $\Delta E$ will be negative. The bomb calorimeter is designed to measure $Q_v$ and thereby to give us a way of determining $\Delta E$ for reactions.

***Processes at constant pressure.*** Chemical and biochemical reactions are much more likely to be conducted at constant pressure (usually 1 atm) than they are at constant volume. For this reason, chemists tend to use the **enthalpy *H*** more often than the internal energy $E$.

$$H = E + PV \quad (6\text{-}3)$$

It follows from Eq. 6-3 that *if the pressure is constant*, $\Delta H_p$ is equal to $\Delta E_p + P\,\Delta V$. Since in a process at constant pressure, $P\,\Delta V$ is exactly the pressure–volume work done on the surroundings, the heat absorbed at constant pressure ($Q_p$) is a measure of $\Delta H_p$.

$$Q_P = \Delta E_P + P\,\Delta V = \Delta H_P \quad (6\text{-}4)$$

Since enthalpy changes can be obtained directly from measurement of heat absorption at constant pressure, even small values of $\Delta H$ for chemical and biochemical reactions can be measured using a microcalorimeter.[11,12] Using the technique of pulsed acoustic calorimetry, changes during biochemical processes can be followed on a timescale of fractions of a millisecond. An example is the laser-induced dissociation of a carbon monoxide–myoglobin complex.[13]

The term *enthalpy* was coined to distinguish $H$ from $E$, but we sometimes tend to be careless about language and many discussions of energy in the literature are in fact about enthalpy. The difference is often not significant because if the pressure–volume work is negligible, $E$ and $H$ are the same.

***Enthalpies of combustion and physiological fuel values.*** The **heat of combustion** ($-\Delta H_c$) of an organic substance is usually determined from $\Delta E_c$, which is measured in a bomb calorimeter. Since $\Delta E_v$ and $\Delta E_p$ are nearly identical, it follows that $\Delta H_p = \Delta E_v + P\,\Delta V$. Here $\Delta V$ is the volume change which would have occurred if the reaction were carried out at constant pressure $P$; thus, $\Delta H_p$ can be estimated by calculation. Since $\Delta H$ is desired for combustion to carbon dioxide, water, elemental nitrogen ($N_2$), and sulfur, correction must be made for the amounts of the latter elements converted into oxides. By these procedures, it has been possible to obtain highly accurate values of $\Delta H_c$ both for biochemical compounds and for mixed foodstuffs. In nutrition, $-\Delta H_c$ is sometimes referred to as the **gross energy**. Values are usually expressed in kilocalories (kcal) by chemists but often as **Cal** (with a capital C) in the nutritional literature.

**Caloric values** of foods (physiological fuel values) are enthalpies of combustion but with an opposite sign, ($-\Delta H_c$), and corrected for energy lost in urine (e.g., as urea) and feces. While enthalpies of combustion of foods are all negative, the caloric values are given as positive numbers. Caloric values for proteins are

**TABLE 6-1**
**Units of Energy and Work and the Values of Some Physical Constants**

| |
|---|
| The joule, SI unit of energy<br>1 J = 1 kg $m^2$ $s^{-2}$<br>= 1 N m (newton meter)<br>= 1 W s (watt second)<br>= 1 C V (coulomb volt) |
| Thermochemical calorie<br>1 cal = 4.184 J |
| Large calorie<br>1 Cal = 1 kcal = 4.184 kJ |
| Work required to raise 1 kg 1 m on earth (at sea level) = 9.807 J |
| Gibbs energy of hydrolysis of 1 mole of ATP at pH 7, millimolar concentrations = −12.48 kcal = −52.2 kJ |
| Work required to concentrate 1 mole of a substance 1000-fold, e.g., from $10^{-6}$ to $10^{-3}$ M = 4.09 kcal = 17.1 kJ |
| Avogadro's number, the number of particles in a mole<br>$N$ = 6.0220 x $10^{23}$ |
| Faraday 1 F = 96,485 C $mol^{-1}$ (coulombs per mole) |
| Coulomb 1 C = 1 A s (ampere second)<br>= 6.241 x $10^{18}$ electronic charges |
| The Boltzmann constant<br>$k_B$ = 1.3807 x $10^{-23}$ J $deg^{-1}$ |
| The gas constant, $R$ = N $k_B$<br>$R$ = 8.3144 J $deg^{-1}$ $mol^{-1}$<br>= 1.9872 cal $deg^{-1}$ $mol^{-1}$<br>= 0.08206 1 atm $deg^{-1}$ $mol^{-1}$<br>and at 25°C $RT$ = 2.479 kJ $mol^{-1}$ |
| The unit of temperature is the kelvin (K); 0°C = 273.16 K |
| ln $x$ = 2.3026 log $x$ |
| One atmosphere (atm) = 101.325 kilopascals (kPa) |

calculated for the conversion of the nitrogen to urea, the major nitrogenous excretion product in mammals, rather than to elemental nitrogen. Typical values are shown in Table 6-2.

**TABLE 6-2**
**Caloric Values of Food Components**

| Component | Caloric values per gram | |
|---|---|---|
| Carbohydrates | 4.1 kcal | 17 kJ |
| Pure glucose | 3.75 kcal | 15.7 kJ |
| Lipids | 9.3 kcal | 39 kJ |
| Proteins[a] | 4.1 kcal | 17 kJ |

[a] Nitrogen excreted as urea.

From a thermochemical viewpoint, can a human or animal be regarded as just a catalyst for the combustion of foodstuffs? To answer this question, large calorimeters were constructed into which an animal or a human being was placed. If, while in the calorimeter, the subject neither gained nor lost weight, the heat evolved should have been just equal to $-\Delta H$ for combustion of the food consumed to $CO_2$, water, and urea. That this prediction was verifed experimentally does not seem surprising, but at the time that the experiments were first done in the early years of the century there may have been those who doubted that the first law of thermodynamics applied to mammals.

In practice, animal calorimetry is quite complicated because of the inherent difficulty of accurate heat measurements, uncertainties about the amount of food stored, and the necessity of corrections for $\Delta H_c$ of the waste products. However, the measurement of energy metabolism has been of considerable importance in nutrition and medicine. Indirect methods of calorimetery have been developed for use in measuring the **basal metabolic rate** of humans. For a good discussion see White *et al*.[14]

The basal metabolic rate is the rate of heat evolution in the resting, postabsorptive state, in which the subject has not eaten recently. In this condition, stored foods provide the energy and are oxidized at a relatively constant rate. The basal metabolic rate tends to be *proportional to the surface area*; which can be approximated (in units of $m^{-2}$) as 1/60 [height (cm) × mass (kg)].[15] For a young adult female, the basal metabolic rate is typically ~154 kJ $h^{-1}$ $m^{-2}$ and for a young adult male ~172 kJ $h^{-1}$ $m^{-2}$. This is ~320 – 360 kJ $h^{-1}$ for a 70-kg person. Note that 360 kJ $h^{-1}$ is the same as the power output of a 100-watt lightbulb. While there is considerable variation among individuals, basal metabolic rates far below or above normal may indicate a pathological condition such as an insufficiency or oversupply of the thyroid hormone thyroxine. Metabolic rates fall somewhat below the basal value during sleep and are much higher than basal during hard exercise. A human may attain rates as high as 2500 kJ (600 kcal) per hour. At a basal rate of 320 kJ (76 kcal) per hour, a person requires 7680 kJ (1835 kcal) each 24 h to supply his or her basal needs, plus additional energy during periods of muscular exercise. Routine light exercise as in the office or during housework increases metabolism to about double the basal rate. Although the caloric values in Table 6-2 are reliable for prediction of metabolic energy needs, they must be adjusted to predict the efficiency of utilization for growth. In one study[16] a group of rats deposited 28% of the available energy from sucrose as body protein and fats, but fats were deposited with an efficiency of 36%.

## 3. The Second Law of Thermodynamics

Why does heat flow from a warm body into a cold one? Why doesn't it ever flow in the reverse direction? We can see that differences in temperature control the direction of flow of heat, but this observation raises still another question: What *is* temperature? Reflection on these questions, and on the interconversion of heat and work, led to the discovery of the second law of thermodynamics and to the definition of a new thermodynamic function, the **entropy *S***.

Consider the melting of ice. This is a phase transition that usually takes place at constant temperature and pressure. At a temperature just above 0°C ice melts completely, but at a temperature just below 0°C it does not melt at all. At 0°C we have an equilibrium. In the language of thermodynamics, the melting of ice at 0°C is a **reversible reaction**. What criterion could be used to predict this behavior for water? For many familiar phenomena, e.g., combustion, a spontaneous reaction is accompanied by the evolution of a large amount of heat, i.e., $\Delta H$ is negative. However, when ice melts it *absorbs* heat. The $\Delta H$ of fusion amounts to 6.008 kJ mol$^{-1}$ at 0°C and is nearly the same just below 0°C, where the ice does not melt, and just above 0°C, where the ice melts completely. In the latter case, the melting of ice is a spontaneous reaction for which $\Delta H$ is positive. It is clear from such facts that *the sign of the enthalpy change does not serve as a criterion of spontaneity*.

A correct understanding of the ice–water transition came when it was recognized that when ice melts not only does $H$ increase by 6.008 kJ mol$^{-1}$, as the molecules acquire additional internal energy of translation, vibration, and rotation, but also *the molecules become more disordered*. Although historically entropy was introduced in a different context, it is now recognized to be a measure of "microscopic disorder." When ice melts, the entropy $S$ increases because the structure becomes less ordered.

The second law of thermodynamics is stated in many different ways, but the usual mathematical formulation asserts that for the universe (or for an isolated system)

$$\Delta S \text{ (system + surroundings)} = 0 \quad \text{for reversible processes}$$

$$\Delta S > 0 \quad \text{for real (nonreversible) processes} \tag{6-5}$$

The second law is sometimes stated in another way: *The entropy of the universe always increases*.

The second law also defines both $S$ and the thermodynamic temperature scale as follows:

$$dS_{\text{reversible}} = q/T \tag{6-6}$$

Here $q$ is an infinitesimal quantity of heat absorbed from the surroundings by the system and $T$ is measured in kelvins (K). For a *reversible phase transition* such as the melting of ice at constant pressure and temperature, the change in entropy of the $H_2O$ is just $\Delta H/T$.

$$\Delta S)_{P,T,\text{reversible}} = Q/T = \Delta H/T \tag{6-7}$$

Entropy is measured in units of joules per kelvin (or °C) or calories per K, the latter sometimes being abbreviated as e. u. (entropy units). Since the melting of ice at 0°C is a reversible process, the second law asserts that the entropy of the surroundings decreases by the same amount that the entropy of the water increases. The value of $T\,\Delta S$ is numerically equal to the heat of fusion, 6.008 kJ mol$^{-1}$ in the case of water at 0°C. Thus, the entropy increase in the ice as it melts at 0°C is 6.008 x $10^3$ J/273.16 K = 22.0 J K$^{-1}$.

***The thermodynamic temperature.*** The definition of thermodynamic temperature in kelvins (Eq. 6-8) also follows from Eq. 6-6. See textbooks of thermodynamics for further treatment.

$$T = (\partial E/\partial S)_V = (\partial H/\partial S)_P \tag{6-8}$$

The entropy of a substance can be given a precise mathematical formulation involving the degree of molecular disorder (Eq. 6-9).

$$S = k_B \ln \Omega \tag{6-9}$$

Here $k_B$ is the **Boltzmann constant** (see Table 6-1) and $\Omega$ is given precisely as the number of microscopic states (different arrangements of the particles) of the system corresponding to a given macroscopic state, i.e., to a given temperature, pressure, and quantity. It increases as volume or temperature is increased and in going from solid to liquid to gaseous states. Equation 6-9 is not part of classical thermodynamics (which deals only with macroscopic systems, i.e., with large collections of molecules). However, using the methods of statistical thermodynamics,[17] this equation can be used to predict the entropies of gases.

The *racemization of an amino acid* provides a biochemical example that can be related directly to Eq. 6-9. A solution of an L-amino acid will be efficiently changed into the racemic mixture of 50% D and 50% L by the action of an enzyme (a **racemase**) with no uptake or evolution of heat. Thus, $\Delta H = 0$ and the only change is an entropy change. Let us designate $\Omega$ for the pure isomer as $\Omega'$. Since there are just two choices of configuration for each of the N molecules in 1 mole of the racemate we see that for the racemate

$$\Omega = 2^N \Omega' \tag{6-10}$$

Applying Eq. 6-9 we calculate $\Delta S$ as follows:

$$\Delta S = k_B (\ln 2^N + \ln \Omega') - k_B \ln \Omega'$$
$$= Nk_B \ln 2 = R \ln 2 = 5.76 \text{ J K}^{-1} \text{ mol}^{-1} \quad (6\text{-}11)$$

***Entropies from measurement of heat capacities.*** It follows from Eq. 6-9 that $S = 0$ when $T = 0$ for a perfect crystalline substance in which no molecular disorder exists. The *third law of thermodynamics* asserts that as the thermodynamic temperature $T$ approaches 0 K the entropy $S$ also approaches zero for perfect crystalline substances. From this it follows that at any temperature above 0 K, the entropy is given by Eq. 6-12.

$$S = \int_0^T C_p \, d \ln T \quad (6\text{-}12)$$

In this equation $C_p$ is the heat capacity at constant pressure:

$$C_P = (\partial H / \partial T)_P \quad (6\text{-}13)$$

If $C_p$ is measured at a series of low temperatures down to near zero K, Eq. 6-12 can be used to evaluate the absolute entropy $S$. If phase transitions occur as the temperature is raised, entropy increments given by Eq. 6-7 must be added to the value of $S$ given by Eq. 6-12. For a few compounds, such as water (Chapter 2), molecular disorder is present in the crystalline state even at 0 K. For these substances a term representing the entropy at 0 K must be added to Eq. 6-12.

The entropies of a few substances are given in Table 6-3. Notice how the entropy increases with increasing complexity of structure, with transitions from solid to liquid to gas, and with decreasing hardness of solid substances.

Measurements of $C_p$ versus temperature for solutions of macromolecules or for biological membranes (Fig. 8-9) over a narrower temperature range are also of interest. These can be obtained with a **differential scanning calorimeter**[18,19] or by an indirect procedure.[20] Denaturation of polymers or phase changes in membranes may be observed. Larger values of $C_p$ are observed for open, denatured, or random-coil structures that are usually present at higher temperatures than for tightly folded molecules.

**TABLE 6-3**
**Entropies of Selected Substances[a]**

| Substance | State[b] | Entropy $S$ | |
|---|---|---|---|
| | | cal $K^{-1}$ $mol^{-1}$ | J $K^{-1}$ $mol^{-1}$ |
| C (diamond) | s | 0.55 | 2.3 |
| C (graphite) | s | 1.36 | 5.7 |
| Cu | s | 8.0 | 33 |
| Na | s | 12.2 | 51 |
| $H_2O$ (ice) | s | 9.8 | 41 |
| $H_2O$ | l | 16.7 | 70 |
| $H_2O$ | g (1 atm) | 45.1 | 189 |
| He | g | 30.1 | 126 |
| $H_2$ | g | 31.2 | 131 |
| $N_2$ | g | 45.8 | 192 |
| $CO_2$ | g | 51.1 | 214 |
| Benzene | g | 64.3 | 269 |
| Cyclohexane | g | 71.3 | 298 |

[a] All values are given in entropy units (e.u.) of calories per Kelvin per mole and in joules per Kelvin per mole at 25°C (298.16 K).
[b] Here s stands for solid, l for liquid, and g for gaseous.

## 4. A Criterion of Spontaneity: The Gibbs Energy

We have seen that while many spontaneous processes, e.g., combustion of organic compounds, are accompanied by liberation of heat (negative $\Delta H$), others are accompanied by absorption of heat from the surroundings (positive $\Delta H$). An example of the latter is the melting of ice at a temperature just above 0°C, during which there is a large increase in the entropy of the water. As we have seen, at 0°C at equilibrium $T\,\Delta S$ is just equal to $-\Delta H$ (Eq. 6-7).

The recognition that $\Delta H - T\,\Delta S = 0$ for a system at equilibrium led J. W. Gibbs to realize that the proper thermodynamic function for determining the spontaneity of a reaction is what is now known as the **Gibbs energy** or Gibbs function ***G*** (Eq. 6-14).

$$G = H - TS \quad (6\text{-}14)$$

In the older literature the Gibbs energy was usually called the **free energy** or Gibbs free energy and was often given the symbol $F$. For a process at constant temperature and pressure the change in $G$ is given by Eq. 6-15 in which all quantities refer to the system.

$$\Delta G_{T,P} = \Delta H - T\,\Delta S \quad (6\text{-}15)$$

For a reversible (equilibrium) process doing only pressure–volume work:

$$\Delta G)_{T,\text{reversible}} = \Delta H - T\,\Delta S = 0 \quad (6\text{-}16)$$

It can also be shown readily that $\Delta G$ *is negative for any spontaneous* (irreversible) *process*. Such a process is called **exergonic**. Likewise, if $\Delta G$ is positive, a given

reaction will *not* proceed spontaneously and is called **endergonic**. The magnitude of the decrease in the Gibbs energy ($-\Delta G$) is a direct measure of the maximum work which could be obtained from a given chemical reaction if that reaction could be coupled in some fashion reversibly to a system able to do work. It represents the maximum amount of electrical work that could be extracted or the maximum amount of muscular work or osmotic work obtainable from a reaction in a biological system. In any real system, the amount of work obtainable is necessarily less than $-\Delta G$ because real processes are irreversible, i.e., entropy is created.

Returning to the older assumption that the magnitude of $\Delta H$ might be an index of work obtainable, we note that $T\,\Delta S$ amounts to only a few kilojoules for most reactions. Therefore, if $\Delta H$ is large, as in the combustion of foodstuffs, it is not greatly different from $\Delta G$ for the same process. Therefore, we can justify use of the caloric value of a food as an approximate measure of the work obtainable from its metabolism in the body.

## 5. Practical Thermochemistry

For thermodynamic data to be useful in chemical calculations, we must agree upon **standard states** for elements and compounds. If we wish to talk about the change in the Gibbs energy that occurs when one or more pure compounds are converted to other pure substances, we must agree upon a state (crystalline, liquid, gaseous, or in solution) and upon a pressure (especially when gases are involved) at which the data apply. The standard pressure is usually 1 atm. Standard states of the elements are the pure *crystalline*, *solid*, or *gaseous materials*, e.g., C (graphite), S (crystalline, rhombic), P (crystalline, white), and $N_2$, $O_2$, and $H_2$ (gaseous). It is also essential to specify the temperature. Thermodynamic data are most often given for 25°C, but there is a standard state for each substance at each temperature.

It is usually impractical to measure the values of $G$ or $H$, but $\Delta G$ and $\Delta H$ for a chemical reaction can be evaluated. Changes in Gibbs energy can be calculated from tables of $\Delta G$ for formation of compounds from the elements (Eq. 6-17). These values of $\Delta G_f$ can be obtained experimentally by measuring $\Delta H$ of combustion for the compound of interest and for $H_2$, elemental carbon, and other elements present in the compound and also by obtaining entropies from heat capacity measurements. Many other tabulated $\Delta G_f$ values have been obtained indirectly utilizing data from equilibrium constants. The resulting **standard Gibbs energies of formation** are given the symbol $\Delta G_f°$. *The values of $\Delta G_f°$ for the elements in their standard states are all exactly zero*.

***Summing changes in Gibbs energy.*** A convenient feature of thermodynamic calculations is that if two or more chemical equations are summed, $\Delta G$ for the resulting overall equation is just the sum of the $\Delta G$'s for the individual equations as illustrated in Eqs. 6-17 to 6-20. The same applies for $\Delta H$ and $\Delta S$.

$$\mathrm{CH_3COOH\ (l) \longrightarrow 2\,C + 2\,H_2 + O_2}$$
$$-\Delta G_f° \text{ for acetic acid} = +396.4 \text{ kJ mol}^{-1} \quad (6\text{-}17)$$

$$\mathrm{2\,O_2 + 2\,C \longrightarrow 2\,CO_2}$$
$$2\,\Delta G_f° \text{ for } CO_2 = -788.8 \text{ kJ mol}^{-1} \quad (6\text{-}18)$$

$$\mathrm{O_2 + 2\,H_2 \longrightarrow 2\,H_2O}$$
$$2\,\Delta G_f° \text{ for } H_2O = -474.4 \text{ kJ mol}^{-1} \quad (6\text{-}19)$$

$$\overline{\mathrm{CH_3COOH\ (l) + 2\,O_2 \longrightarrow 2\,CO_2 + 2\,H_2O\ (l)}}$$
$$\Delta G_c° \text{ for acetic acid} = -866.8 \text{ kJ mol}^{-1} \quad (6\text{-}20)$$

In this example an equation for the decomposition of acetic acid into its elements (Eq. 6-17) has been summed with Eqs. 6-18 and 6-19, which represent the formation of the proper number of molecules of $CO_2$ and $H_2O$ from the elements. The sum of the three equations gives the equation for the combustion of acetic acid to $CO_2$ and water, and the sum of the $\Delta G$ values for the three equations gives $\Delta G$ for combustion of acetic acid. The resulting value of $\Delta G$ is for combustion of pure liquid acetic acid by oxygen at 1 atm to give $CO_2$ at 1 atm and pure liquid water, all reactants and products being in their standard states.

The process described in the preceding paragraph is represented by Eq. 6-21, which is a general equation for calculation of $\Delta G°$ for any reaction from $\Delta G_f°$ of products and reactants.

$$\Delta G° = \Sigma \Delta G_f° \text{ (products)} - \Sigma \Delta G_f° \text{ (reactants)} \quad (6\text{-}21)$$

How does the change in Gibbs energy vary if we go from the standard state of a compound to some other state? Consider a change of pressure in a gas. It is easy to show (see any thermodynamics text) that

$$(\partial G/\partial P)_T = V \quad (6\text{-}22)$$

Using Eq. 6-22 together with the perfect gas law, we obtain the relationship (Eq. 6-23) between the Gibbs energy $\overline{G}$ of one mole of a substance at pressure $P$ and the standard Gibbs energy $\overline{G}°$ at pressure $P°$.

$$\overline{G} - \overline{G}° = RT \ln \frac{P}{P°} = RT\ \Delta \ln P \quad (6\text{-}23)$$

Here the bar over the symbol $G$ indicates that the Gibbs energy is for one mole of substance. Since $P°$ is by definition 1 atm, the Gibbs energy change per mole upon changing the pressure from $P°$ to $P$ is just $RT \ln P$.

***Reactions in solution.*** It is customary in books on thermodynamics to develop most of the important thermodynamic equations as applied to a perfect gas, but we will move at this point to a consideration of biochemical substances in solution. Biochemists are usually interested in the behavior of substances dissolved in relatively dilute aqueous solutions but also in cytoplasm, in which some concentrations may be very high. Sometimes the interest may be in nonaqueous solutions. In any case, it is necessary to establish a standard state for the solute. The standard state of a substance in aqueous solution is customarily taken as a strictly hypothetical one **molal** solution (one mole of solute per kilogram of water) *whose properties are those of a solute at infinite dilution*. An equation exactly analogous to Eq. 6-23 can be written relating the Gibbs energy of one mole of dissolved solute $\overline{G}_i$ to the Gibbs energy $\overline{G}_i°$ in the hypothetical standard state of unit activity and to the **activity** $a_i$ of the solute (Eq. 6-24).

$$\overline{G}_i = \overline{G}_i° + RT \ln a_i \tag{6-24}$$

Here the subscript $i$ designates a particular component in a solution which also contains solvent and, perhaps, other components. To be precise, $\overline{G}_i$ is a *partial molar Gibbs energy*, i.e., the changes in total Gibbs energy of a very large volume of solution when one mole of the component is added.

From Eq. 6-24 it follows that the *Gibbs energy change for dilution* from one activity $a_1$ to another $a_2$ is:

$$\Delta\overline{G} \text{ (dilution from } a_1 \text{ to } a_2) = RT \ln (a_2/a_1) \tag{6-25}$$

Equation 6-24 and the equations that follow from it apply to molal activities. However, the concentration can be substituted for activity in very dilute solution where the behavior of the dissolved molecules approximates that of the hypothetical ideal solution for which the standard state is defined. For any real solution, the activity can be expressed as the product of an activity coefficient and the concentration (Eq. 6-26).

$$a = \gamma c \tag{6-26}$$

where $a$ = activity, $\gamma$ = activity coefficient, and $c$ = molal concentration. Thus, to use the tabulations of thermodynamic functions for substances in solution to predict behavior in other than very dilute solution, we must multiply the concentration of every component by the appropriate activity coefficient. For the approximate calculations which are often of interest to biochemists, it is customary to equate concentration with activity. Furthermore, in dilute solutions, the more usual **molar** concentrations (moles per liter) are nearly equal to molal concentrations. The same equations can be used with *mole fractions* rather than molal concentrations.

***ΔG° and the equilibrium constant.*** Consider the following generalized chemical equation (Eq. 6-27) for reaction of $a$ moles of A with $b$ moles of B to give products C and D, etc.

$$a\,\mathrm{A} + b\,\mathrm{B} + \cdots = c\,\mathrm{C} + d\,\mathrm{D} \cdots \tag{6-27}$$

The standard Gibbs energy change $\Delta G°$ for the process is given by Eq. 6-28:

$$\Delta G° = \mathrm{c}\,\overline{G}°(\mathrm{C}) + \mathrm{d}\,\overline{G}°(\mathrm{D}) + \cdots - \mathrm{a}\,\overline{G}°(\mathrm{A}) - \mathrm{b}\,\overline{G}°(\mathrm{B}) \cdots \tag{6-28}$$

The symbol $G°$ (A) designates the Gibbs energy of A, etc. The value of $\Delta G$ for any desired concentrations of reactants or products can be related to this $\Delta G°$ by applying to each component Eq. 6-24 with the following result:

$$\Delta G = \Delta G° + RT \ln \frac{a_C{}^c\, a_D{}^d \cdots}{a_A{}^a\, a_B{}^b \cdots} \tag{6-29}$$

Here $a_C$ represents the activity of component C, etc. This useful equation permits us to calculate $\Delta G$ for the low concentrations usually found in biochemical systems. These are more often in the millimolar range or less rather than approaching the hypothetical 1 M of the standard state. Often concentrations are substituted in Eq. 6-29 for activities:

$$\Delta G \approx \Delta G° + RT \ln \frac{[\mathrm{C}]^c[\mathrm{D}]^d}{[\mathrm{A}]^a[\mathrm{B}]^b} \tag{6-30}$$

Equation 6-29 is used in another way by noting that $\Delta G = 0$ when a system is at equilibrium and that at equilibrium the product $a_C{}^c\, a_D{}^d \ldots / a_A{}^a\, a_B{}^b \ldots$ is just the equilibrium constant $K$. It follows that

$$\begin{aligned} \Delta G° &= -RT \ln K = -2.303RT \log K \\ &= -19.145T \log K \text{ J mol}^{-1} \\ &= -5.708 \log K \text{ kJ mol}^{-1} \text{ at } 25°\mathrm{C} \\ &= -1.364 \log K \text{ kcal mol}^{-1} \text{ at } 25°\mathrm{C} \end{aligned} \tag{6-31}$$

Although the units of $\Delta G°$ are kJ mol$^{-1}$, the Gibbs energy change in Eq. 6-31 is that for the reactions of $a$ moles of

A, *b* moles of B, etc., as in the equation used to define *K* (Eq. 6-27 in this instance). It is also important to realize that the log term in Eq. 6-31 must be unitless. Although we usually write ln *K* or log *K*, *K* here represents *K* / *Q*, where *Q* = 1 because it has the same form as *K* but with all components in their standard states. Since the units of *K* and *Q* are the same, log *K* is unitless. Similar considerations apply to expressions of *K* in exponential form.

***Activity coefficients and concentration equilibrium constants.*** Strictly speaking, Eq. 6-31 applies only to thermodynamic equilibrium constants –that is, to constants that employ activities rather than concentrations. The experimental determination of such constants requires measurements of the apparent equilibrium constant or **concentration equilibrium constant**[21] $K_c$ at a series of different concentrations and extrapolation to infinite dilution (Eq. 6-32).

$$K_c = \text{concentration equilibrium constant} = \frac{[C]^c[D]^d}{[A]^a[B]^b} \text{ at equilibrium} \qquad (6\text{-}32)$$

Extrapolation of $K_c$ to infinite dilution to give *K* is usually easy because the activity coefficients of most ionic substances vary in a regular manner with **ionic strength** and follow the **Debye–Hückel** equation (Eq. 6-33) in very dilute solutions (ionic strength < 0.01).

$$\log \gamma = -0.509 z_1 z_2 \sqrt{\mu} \qquad (6\text{-}33)$$

The integers $z_1$ and $z_2$ are the numbers of charges (valences) for the cation and anion of the salt. The ionic strength (μ, or I) is evaluated as follows:

$$\mu = \tfrac{1}{2} \Sigma_i c_i z_i^2 \qquad (6\text{-}34)$$

Here $c_i$ are the molar concentrations of the ions. The summation is carried out over all the ions present. The activity coefficient γ (Eq. 6-33) is the mean activity coefficient for both the cation and anion.

Equation 6-33 suggests that extrapolation of equilibrium constants to infinite dilution is done appropriately by plotting log $K_c$ vs $\sqrt{\mu}$. For example, Fig. 6-1 shows plots of $pK'_a$ for dissociation of $H_2PO_4^-$, $AMP^-$, and $ADP^{2-}$, and $ATP^{3-}$ vs $\sqrt{\mu}$. The variation of $pK'_a$ with $\sqrt{\mu}$ at low concentrations (Eq. 6-35) is derived by application of the Debye-Hückel equation (Eq. 6-33):

$$pK_a' = pK_a - 0.509 (z_A^2 - z_{HA}^2) \sqrt{\mu} \qquad (6\text{-}35)$$

Straight lines of slope $-0.509 (z_A^2 - z_{HA}^2)$ are expected. The observed (negative) slopes (Fig. 6-1) are ~1.5 for $H_2PO_4^-$ and $AMP^-$, ~2.5 for $ADP^{2-}$, and ~3.5 for $ATP^{3-}$. The data over the entire range of ionic strength are fitted by empirical relationships of the type of Eq. 6-35a:

$$pK_a' = pK_a - a\sqrt{\mu} + b\mu \quad \text{for } \mu < 0.2 \qquad (6\text{-}35a)$$

in which *a* and *b* are empirically determined constants. For example, for $H_2PO_4^-$ $a = 1.52$ and $b = 1.96$. The value of $pK_a$ found was 7.18, about 0.22 greater than the value at μ = 0.2, an ionic strength more commonly used in the laboratory and close to that found in tissues. Note that the difference between the extrapolated $pK_a$ for $ATP^{3-}$ of 7.68 and the observed value of ~ 7.04 at μ = 0.2 is even greater. Serious errors can be introduced into calculations by using extrapolated values for *K* for solutions of appreciable ionic strength. The errors will be maximal for ions of high charge type such as $ATP^{3-}$ and $ATP^{4-}$.

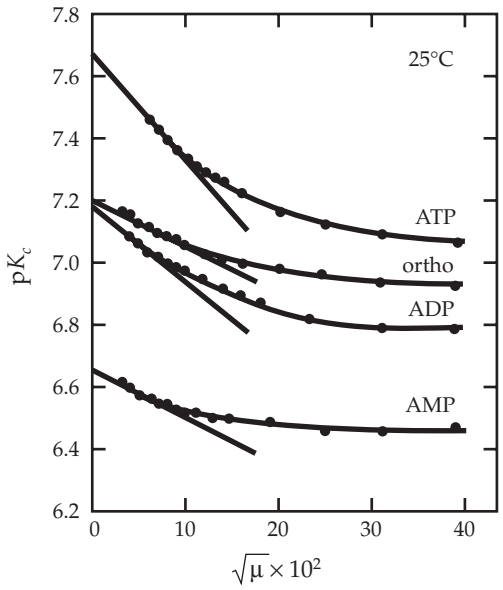



**Figure 6-1** The apparent $pK_a$ values for secondary ionizations of AMP, ADP, ATP, and $H_3PO_4$ (abbreviated ortho) plotted against $\sqrt{\mu}$. Temperature: 25°C. From R. C. Phillips *et al.*[22]

Another problem with equilibrium constants for reactions that use or produce hydrogen ions is that there is no rigorous relationship between pH and $a_{H^+}$ or $[H^+]$. Indeed, the concept of an activity of a single ion has little meaning in thermodynamics. Nevertheless, in the pH range of interest to biochemists, results that are very close to those obtained by more rigorous methods are achieved by assuming that the pH meter

responds to hydrogen ion activity. The almost universal practice of biochemists is to assume the pH meter reading obtained with a glass electrode equal to – log $a_{H^+}$ and to substitute the value of $a_{H^+}$ so obtained for $[H^+]$ in defining the concentration equilibrium constant, $K_c$. It is also customary in most branches of chemistry to use values of equilibrium constants and of Gibbs energy which have *not* been extrapolated to $\mu = 0$. Thus most values of $K$ and $\Delta G$, including many of those in this book, are actually of $K_c$ and $\Delta G_c$. An international Commission on Biothermodynamics[21] recommended that values of $K_c$ be measured with the lowest effective buffer concentration and that the ionic strength be brought to 0.1 with KCl.

***Changes in equilibria with temperature.*** At constant pressure $\Delta G$ varies with absolute temperature as follows:

$$\frac{d(\Delta G/T)}{dT} = -\Delta H/T^2 \tag{6-36}$$

The corresponding variation in $K$ is described by the **van't Hoff equation**:

$$\frac{d \ln K}{dT} = \frac{\Delta H^\circ}{RT^2}$$

or

$$\frac{d \ln K}{d(1/T)} = \frac{-\Delta H^\circ}{R}$$

or

$$\Delta H^\circ(\text{kJ mol}^{-1}) = \frac{-0.01914\ d \log_{10} K}{d(1/T)} \tag{6-37}$$

If $\Delta H^\circ$ can be assumed constant over the temperature range of an experiment, a plot of $\ln K$ vs $1/T$ provides a convenient estimate of $\Delta H^\circ$ (or $\Delta H'$ if $\ln K'$ is plotted). The slope of the line will be $-\Delta H^\circ/R$. Since $\Delta G^\circ$ can be calculated from $K$, the method also permits evaluation of $\Delta S^\circ$ using Eq. 6-15. However, unless great care is taken the method is of low accuracy[23] and it is preferable to establish $\Delta H$ by direct calorimetry. Also, especially for proteins, the assumption that $\Delta H^\circ$ is constant over a significant temperature range may be erroneous.

From observations at only two temperatures, $T_1$, and $T_2$, Eq. 6-37 becomes

$$\Delta H^\circ = R \ln (K_2/K_1)[T_1T_2/(T_2 - T_1)] \text{ kJ mol}^{-1} \tag{6-37a}$$

## 6. Thermodynamics and Life Processes

Can thermodynamics be applied to living organisms? Classical thermodynamics deals with equilibria, but living beings are never in equilibrium. The laws of thermodynamics are usually described as statistical laws. How can such laws apply to living things, some of which contain all of their genetic information in a single molecule of DNA?[24,25] The ideal, reversible reactions of classical thermodynamics occur at infinitesimal speeds. How can thermodynamics be applied to the very rapid chemical reactions that take place in organisms? One answer is that thermodynamics can be used to decide *whether or not a reaction is possible* under given conditions. Thus, if we know the steady-state concentrations of reactants and products within a cell, we can state whether a reaction will or will not tend to go in a given direction.

We may still ask whether there are generalities comparable to the laws of thermodynamics that apply to the kind of steady state or "dynamic equilibrium" that exists in organisms. Lars Onsager showed that such relationships can be found for conditions that are *near equilibrium*. Ilya Prigogine and associates extended Onsager's findings and showed that under conditions that are *far from equilibrium* the system tends to become unstable and to spontaneously develop new structures, which Prigogine calls **dissipative structures**.[26–28] Vortices in flowing water, tornados and hurricanes are examples of dissipative structures. The maintenance of dissipative structures depends upon a flow of energy and matter through the system. The flow of energy is sufficient to "organize" a system. Thus, if a flask of water is placed on a hot plate, a cycle is established. The water moves via cyclic convection currents that develop as a result of the flow of energy through the system. Morowitz reckoned that the $6 \pm 3 \times 10^{18}$ kJ/year of solar energy that falls on the earth supplies the organizing principle for life.[28] Just as it drives the great cycles within the atmosphere and within the seas, it gives rise to the branching and interconnecting cycles of metabolism. This idea may even make the spontaneous development of the organized systems that we call life from inanimate precursors through evolution seem a little more understandable.

A characteristic of this nonequilibrium or irreversible thermodynamics is that time is explicitly introduced. Furthermore, open systems, in which materials and energy flow into and out of the system, are considered. Clearly, a living organism is an open system not a closed one of classical thermodynamics. Because of the flow of materials concentration gradients are set up and transport phenomena often become of primary importance. Articles and books that provide an introduction to nonequilibrium thermodynamics and to the literature in the field include the following.[10,26,28–34] Whether these methods can be applied in a practical way to metabolic systems has been debated.[35,36]

## TABLE 6-4
## Gibbs Energies of Formation and of Oxidation at 25°C for Compounds of Biochemical Interest[a,b]

| Compound | Formula | $\Delta G_f^\circ$ (kJ mol$^{-1}$) | $\Delta G_c^\circ$ (kJ mol$^{-1}$) | For oxidation by $NAD^+$ | | Number of electrons |
|---|---|---|---|---|---|---|
| | | | | $\Delta G'^\circ_{ox}$ (kJ mol$^{-1}$) | $\Delta G'^\circ_{ox}$ (pH 7) (kJ mol$^{-1}$) | |
| Acetaldehyde | $C_2H_4O$ | –139.7 | –1123.5 | 171.5 | –28.3 | 10 |
| Acetic acid | $C_2H_4O_2$ | –369.4 | –866.8 | 169.2 | 9.3 | 8 |
| Acetate$^-$ | | –369.2 | –894.0 | 142.0 | 22.1 | 8 |
| Acetyl-CoA | | –374.1* | –889.1* | 146.9* | –13.0* | 8 |
| Acetyl-*P* | | –1218.4 | –901.7 | 134.3 | –25.6 | 8 |
| Acetylene[c] | $C_2H_2$ | 209.2 | –1235.2 | 59.8 | –140.0 | 10 |
| Acetoacetate$^-$ | $C_4H_5O_3^-$ | –493.7 | –1795.4 | 276.5 | –3.2 | 16 |
| Acetone | $C_3H_6O$ | –161.2 | –1733.6 | 338.4 | 18.7 | 16 |
| *cis*-Aconitate$^{3-}$ | $C_6H_3O_6^{3-}$ | –920.9 | –2157.0 | 173.9 | –65.8 | 18 |
| L-Alanine | $C_3H_7O_2N$ | –371.3 | –1642.0 | 300.4 | 0.8 | 15 |
| L-Asparagine | $C_4H_8O_3N_2$ | –526.6 | –1999.7 | 331.2 | –28.4 | 18 |
| L-Aspartate$^-$ | $C_4H_6O_4N^-$ | –700.7 | –1707.0 | 235.4 | –24.3 | 15 |
| *n*-Butanol | $C_4H_{10}O$ | –171.8 | –2591.7 | 516.2 | 36.7 | 24 |
| *n*-Butyric acid | $C_4H_8O_2$ | –380.2 | –2146.1 | 443.8 | 44.2 | 20 |
| *n*-Butyrate$^-$ | $C_4H_7O_2^-$ | –352.6 | –2173.7 | 416.2 | 56.6 | 20 |
| Butyryl-CoA | | –357.5* | –2168.8* | 421.1* | 21.5* | 20 |
| Caproate$^-$ | $C_6H_{11}O_2^-$ | –329.7 | –3459.7 | 684.1 | 84.7 | 32 |
| $CO_2$ (g) | | –394.4 | 0.0 | 0.0 | 0.0 | 0 |
| $CO_2$ (aq) | | –386.2 | –8.2 | –8.2 | –8.2 | 0 |
| $HCO_3^-$ | | –587.1 | –44.5 | –44.5 | –4.6 | 0 |
| CO (g) | | –137.3 | –257.1 | 1.9 | –38.1 | 2 |
| Citrate$^{3-}$ | $C_6H_5O_7^{3-}$ | –1166.6 | –2148.4 | 182.5 | –57.3 | 18 |
| Creatine | $C_4H_9O_2N_3$ | –264.3 | –2380.6 | 338.8 | –80.8 | 21 |
| Creatinine | $C_4H_7ON_3$ | –28.9 | –2378.8 | 340.6 | –79.0 | 21 |
| Crotonate$^-$ | $C_4H_5O_2^-$ | –275.7 | –2013.4 | 317.5 | –2.1 | 18 |
| Cysteine | $C_3H_7O_2NS$ | –339.8 | –2178.3 | 541.1 | 121.5 | 21 |
| Cystine | $C_6H_{12}O_4N_2S_2$ | –665.3 | –4133.8 | 1046.0 | 246.9 | 40 |
| Dihydroxyacetone-*P*[d] | | –1293.2 | –1458.4 | 95.5 | –144.2 | 12 |
| Erythrose 4-*P*[d] | | –1439.1 | –1944.1 | 127.8 | –191.9 | 16 |
| Ethanol | $C_2H_6O$ | –181.5 | –1318.8 | 235.1 | –4.6 | 12 |
| Ethylene (g)[c] | $C_2H_4O$ | 68.1 | –1331.3 | 222.7 | –17.1 | 12 |
| Formaldehyde | $CH_2O$ | –130.5 | –501.0 | 16.9 | –63.0 | 4 |
| Formic acid | $CH_2O_2$ | –356.1 | –275.5 | –16.5 | –56.5 | 2 |
| Formate$^-$ | $CHO_2^-$ | –350.6 | –281.0 | –22.0 | –22.0 | 2 |
| Fructose | $C_6H_{12}O_6$ | –915.4 | –2874.1 | 233.8 | –245.7 | 24 |
| Fructose 6-*P*[d] | | –1758.3 | –2888.1 | 219.8 | –259.7 | 24 |
| Fructose di-*P*[d] | | –2600.8 | –2902.5 | 205.4 | –274.1 | 24 |
| Fumaric acid | $C_4H_4O_4$ | –647.1 | –1404.8 | 149.2 | –90.6 | 12 |
| Fumarate$^-$ | $C_4H_3O_4^-$ | –604.2 | –1447.7 | 106.2 | –93.6 | 12 |
| α-D-Galactose | $C_6H_{12}O_6$ | –923.5 | –2865.9 | 242.0 | –237.5 | 24 |
| α-D-Glucose | $C_6H_{12}O_6$ | –917.2 | –2872.2 | 235.6 | –243.8 | 24 |
| Glucose 6-*P* | | –1760.3 | –2886.0 | 221.8 | –257.6 | 24 |
| L-Glutamate$^-$ | $C_5H_8O_4N^-$ | –696.8 | –2342.5 | 376.9 | –2.7 | 21 |
| L-Glutamine | $C_5H_{10}O_3N_2$ | –524.8 | –2633.1 | 474.8 | –4.7 | 24 |
| 3-*P* glycerate$^{-d}$ | | –1515.7 | –1235.9 | 59.0 | –100.8 | 10 |
| 2-*P* glycerate$^{-d}$ | | –1509.9 | –1241.8 | 53.2 | –106.6 | 10 |
| Glyceraldehyde 3-*P*[d] | | –1285.6 | –1466.0 | 87.9 | –151.8 | 12 |
| Glycerol | $C_3H_8O_3$ | –488.5 | –1643.4 | 169.5 | –110.2 | 14 |
| Glycerol-*P* | | –1336.2 | –1652.6 | 160.3 | –119.3 | 14 |
| Glycine | $C_2H_5O_2N$ | –373.5 | –1008.3 | 157.2 | –22.6 | 9 |
| Glycogen | $C_6H_{10}O_5$ | –665.3 | –2887.0 | 220.9 | –258.6 | 24 |
| Glycolate$^-$ | $C_2H_3O_3^-$ | –523.4 | –739.7 | 37.2 | –42.7 | 6 |
| Glyoxylate$^-$ | $C_2HO_3^-$ | –461.1 | –564.9 | –46.9 | –86.9 | 4 |
| $H_2O$ (l) | | –237.2 | 0.0 | 0.0 | 0.0 | 0 |
| $OH^-$ | | –157.3 | –79.9 | –79.9 | –39.9 | 0 |
| $H^+$ | | 0.0 | 0.0 | 0.0 | | 0 |
| $H_2$ (g) | | 0.0 | –237.2 | 21.8 | –18.2 | 2 |
| $H_2O_2$ | | –136.8 | –100.4 | –359.4 | –319.4 | –2 |
| $H_2S$ | | –27.4 | –714.6 | 321.3 | 161.5 | 8 |
| $HS^-$ | | 12.6 | –754.5 | 281.4 | 161.5 | 8 |
| β-Hydroxybutyric acid | $C_4H_8O_3$ | –531.4 | –1994.9 | 336.0 | –23.6 | 18 |
| β-Hydroxybutyrate$^-$ | $C_4H_7O_3^-$ | –506.3 | –2020.0 | 310.9 | –8.8 | 18 |
| Hydroxypyruvate | $C_3H_4O_4$ | –615.9 | –1041.6 | –5.7 | –165.5 | 8 |

## TABLE 6-4 (continued)

| Compound | Formula | $\Delta G_f^\circ$ (kJ mol$^{-1}$) | $\Delta G_c^\circ$ (kJ mol$^{-1}$) | For oxidation by $NAD^+$: $\Delta G^\circ_{ox}$ (kJ mol$^{-1}$) | For oxidation by $NAD^+$: $\Delta G^\circ_{ox}$ (pH 7) (kJ mol$^{-1}$) | Number of electrons |
|---|---|---|---|---|---|---|
| Hypoxanthine | $C_5H_6O$ | 89.5 | –2773.0 | 334.8 | –144.6 | 24 |
| Isocitrate$^{3-}$ | $C_6H_5O_7^{3-}$ | –1160.0 | –2155.1 | 175.8 | –63.9 | 18 |
| α-Ketoglutarate$^{2-}$ | $C_5H_4O_5^{2-}$ | –798.0 | –1885.5 | 186.4 | –53.3 | 16 |
| Lactate$^-$ | $C_3H_5O_3^-$ | –516.6 | –1378.1 | 175.9 | –23.9 | 12 |
| α-Lactose | $C_{12}H_{22}O_{11}$ | –1515.2 | –5826.5 | 389.3 | –569.7 | 48 |
| L-Leucine | $C_6H_{13}O_2N$ | –356.3 | –3551.7 | 721.6 | 62.3 | 33 |
| Mannitol | $C_6H_{14}O_6$ | –942.6 | –3084.0 | 282.8 | –236.6 | 26 |
| Malate$^{2-}$ | $C_4H_4O_5^{2-}$ | –845.1 | –1444.0 | 109.9 | –49.9 | 12 |
| Methane (g) | $CH_4$ | –50.8 | –818.0 | 218.0 | 58.2 | 8 |
| Methanol | $CH_4O$ | –175.2 | –693.5 | 83.4 | –36.4 | 6 |
| $NH_4^+$ | | –79.5 | –276.3 | 112.2 | 12.3 | 3 |
| $NO_2^-$ | | –34.5 | –84.1 | –472.6 | –372.7 | –3 |
| NO (g) | | 86.7 | –86.7 | –345.7 | –305.7 | –2 |
| $NO_3^-$ | | –110.5 | –8.1 | –655.6 | –515.7 | –5 |
| Oxalate$^{2-}$ | $C_2O_4^{2-}$ | –674.9 | –351.1 | –92.1 | –52.1 | 2 |
| Oxaloacetate$^{2-}$ | $C_4H_2O_5^{2-}$ | –797.2 | –1254.7 | 40.2 | –79.7 | 10 |
| $H_3PO_4$ (aq)[e] | | –1147.3 | 0.0 | 0.0 | 0.0 | 0 |
| $H_2PO_4^-$ (aq)[e] | | –1135.1 | –12.1 | –12.1 | 27.8 | 0 |
| $HPO_4^{2-}$ (aq)[e] | | –1094.1 | –53.1 | –53.1 | 26.8 | 0 |
| *n*-Propanol | $C_3H_8O$ | –175.8 | –1956.1 | 374.8 | 15.2 | 18 |
| Isopropanol | $C_3H_8O$ | –185.9 | –1946.0 | 384.9 | 25.3 | 18 |
| Propionate$^-$ | $C_3H_5O_2^-$ | –360.0 | –1534.7 | 278.2 | 38.5 | 14 |
| Pyruvate$^-$ | $C_3H_3O_3^-$ | –474.5 | –1183.1 | 111.9 | –47.9 | 10 |
| Phosphoenolpyruvate$^{3-}$ | | –1269.5 | –1245.0 | 50.0 | –109.8 | 10 |
| Ribose 5-*P*[d] | | –1599.9 | –2414.9 | 175.0 | –224.6 | 20 |
| Ribulose 5-*P*[d] | | –1597.6 | –2417.1 | 172.8 | –226.8 | 20 |
| Sedoheptulose 7-*P*[d] | | –1913.3 | –3364.6 | 261.2 | –298.2 | 28 |
| Sedoheptulose di-*P*[d] | | –2755.8 | –3379.0 | 246.9 | –312.5 | 28 |
| Sorbitol | $C_6H_{14}O_6$ | –942.7 | –3083.9 | 282.9 | –236.5 | 26 |
| Succinate$^{2-}$ | $C_4H_4O_4^{2-}$ | –690.2 | –1598.9 | 214.1 | 14.3 | 14 |
| Succinyl-CoA | | –686.7* | –1602.4* | 210.6* | –29.2* | 14 |
| Sucrose | $C_{12}H_{22}O_{11}$ | –1551.8 | –5789.9 | 425.9 | –533.1 | 48 |
| $SO_4^{2-}$ | | –742.0 | 0.0 | 0.0 | 79.9 | 0 |
| $SO_3^{2-}$ | | –497.1 | –244.9 | 14.1 | 54.0 | 2 |
| $S_2O_3^{2-}$ | | –513.4 | –733.4 | 302.5 | 222.6 | 8 |
| L-Threonine | $C_4H_9O_3N$ | –514.6 | –2130.3 | 330.1 | –49.5 | 19 |
| L-Tyrosine | $C_9H_{11}O_3N$ | –387.2 | –4466.8 | 842.5 | 23.4 | 41 |
| Urea | $CH_4ON_2$ | –203.8 | –664.9 | 112.0 | –7.8 | 6 |
| Uric acid | $C_5H_4O_3N_4$ | –356.9 | –2089.4 | 241.5 | –118.1 | 18 |
| L-Valine | $C_5H_{11}O_2N$ | –360.0 | –2916.5 | 579.9 | 40.5 | 27 |
| Xanthine | $C_5H_5O_2N_4$ | –139.3 | –2425.6 | 293.8 | –125.7 | 21 |
| D-Xylulose | $C_5H_{10}O_5$ | –748.1 | –2409.8 | 180.1 | –219.5 | 20 |

[a] The quantities tabulated are $\Delta G_f^\circ$, the standard free energy of formation from the elements; $\Delta G_c^\circ$, the standard free energy of combustion; $\Delta G^\circ_{ox}$, the standard free energy of oxidation by $NAD^+$ to products NADH + $H^+$, $CO_2$, $H_2O$, $N_2$, $HPO_4^{2-}$, and $SO_4^{2-}$; $\Delta G'_{ox}$ (pH 7), the apparent standard free energy change at pH 7. All values are in kJ mol$^{-1}$ at 25°C in aqueous solution unless indicated otherwise. If a compound is designated (g) the values are for the gaseous phase at 1 atm pressure. The number of electrons involved in complete oxidation to $CO_2$, $H_2O$, $N_2$, and $H_2SO_4$ is given in the final column. If this number is negative, the compound must be reduced to obtain the products, e.g., $2\ NO_3^- + 10\ e^- + 12\ H^+ \rightarrow N_2 + 6\ H_20$. The data for phosphate esters refer to the compounds with completely dissociated phosphate groups (–O–$PO_3^{2-}$). The values of $\Delta G_f^\circ$ for many of these compounds were calculated as $\Delta G_f^\circ$ (nonphosphorylated compound) – $\Delta G^\circ$ for hydrolysis (to $HPO_4^{2-}$, Table 6-6) $-\Delta G_f^\circ$ for $H_2O$ (one molecule for each phosphate ester formed) + $\Delta G_f^\circ$ for $HPO_4^{2-}$ (from this table). Data from Bassman and Krause[d] were used directly. For acyl-CoA derivatives CoA (–SH) is treated as an "element," i.e., the values of $\Delta G_f^\circ$ given and designated with an asterisk (*) are for formation from the elements plus free CoA. The values of $\Delta G_c^\circ$ and $\Delta G_{ox}$ are for oxidation to the usual products plus CoA. Values of $\Delta G^\circ$ of hydrolysis (Table 6-6) were used in computing $\Delta G_f^\circ$ for each of these compounds from that of the corresponding alcohol or carboxylate anion. Another source containing an extensive table of Gibbs energy values is Wilhoit, R. C. (1969) in *Biochemical Microcalorimetry* (Brown, H. D. ed.), pp. 305–317. Academic Press, New York

[b] The major source is Long, C., ed. (1961) *Biochemists Handbook*, pp. 90–92. Van Nostrand, Reinhold, Princeton, New Jersey. Most of the values in this collection are from Burton, K. (1957) *Ergeb. Physiol., Biol. Chem. Exp. Pharmakol.* **49**, 275–298

[c] From Stull, D.R., Westrum, E. F., Jr., and Sinke, G. C. (1969) *The Chemical Thermodynamics of Organic Compounds*. Wiley, New York

[d] Bassham, J. A. and Krause, G. H. (1969) *Biochim. Biophys. Acta.* **189**, 207–221

[e] Van Wazer, J. R. (1958) *Phosphorus and Its Compounds*, Vol. I, p. 889. Wiley (Interscience), New York

## B. Tables of $\Delta G°$ Values for Biochemical Compounds

### 1. Gibbs Energies of Formation

Table 6-4 gives, in the first column, standard values of Gibbs energies of formation from the elements $\Delta G_f°$ for a variety of pure solids, gases, and liquids as well as values for substances in solution at the hypothetical 1 M activity. As an example, consider the value of $\Delta G_f°$ for pure liquid acetic acid, –389.1 kJ mol$^{-1}$. The equation for its formation from the elements is:

$$2\,C\,(s) + 2\,H_2\,(g, 1\,atm) + O_2\,(g, 1\,atm) \longrightarrow C_2H_4O_2\,(l)$$
$$\Delta G_f° = -389.1\ kJ\ mol^{-1} \qquad (6\text{-}38)$$

To obtain the Gibbs energy of formation in aqueous solution, we must have solubility data as well as activity coefficients of acetic acid at various concentrations. From these data the change in Gibbs energy for solution of the liquid acetic acid in water to give aqueous acetic acid in the hypothetical 1 molal standard state (Eq. 6-39) can be obtained.

$$\text{Acetic acid (l)} \longrightarrow \text{acetic acid (aq)}$$
$$\Delta G = -7.3\ kJ\ mol^{-1} \qquad (6\text{-}39)$$

Summing Eqs. 6-38 and 6-39 we obtain:

$$2\,C + 2\,H_2 + O_2 \longrightarrow \text{acetic acid (aq)}$$
$$\Delta G_f° = -396.4\ kJ\ mol^{-1} \qquad (6\text{-}40)$$

In many computations it is convenient to have $\Delta G$ values for single ions, e.g., for acetate$^-$. We can obtain $\Delta G_f°$ of acetate$^-$ (aq) from that of acetic acid (aq) by making use of $\Delta G°$ of dissociation (Eq. 6-41).

$$\text{Acetic acid} \longrightarrow H^+ + \text{acetate}^-$$
$$\Delta G° = -5.708 \log K_a = 5.708\ pK_a$$
$$= +27.2\ kJ\ mol^{-1}\ \text{at}\ 25°C \qquad (6\text{-}41)$$

By convention *we define the Gibbs energy of formation of $H^+$ as zero*. Then, by summing Eqs. 6-40 and 6-41 we obtain $\Delta G_f°$ of acetate$^-$ = –369.2 kJ mol$^{-1}$.

### 2. Gibbs Energies of Dissociation of Protons

Table 6-5 gives thermodynamic dissociation constants and values of $\Delta G°$ and $\Delta H°$ for a number of acids of interest in biochemistry. Some of these values were used in obtaining the values of $\Delta G_f°$ for the ions of Table 6-4. The data of Table 6-5 can also be used in evaluation of Gibbs energy changes for reactions of ionic forms not given in Table 6-4.

### 3. Group Transfer Potentials

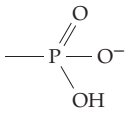
Recall that the equilibria for reactions by which monomers are linked to form biopolymers (whether amides, esters, phosphodiesters, or glycosides) usually favor *hydrolysis* rather than formation (condensation). The equilibrium positions depend on the exact structures. Some linkages are formed easily if monomer concentrations are high enough, but others are never formed in significant concentrations. Likewise, hydrolysis may be partial at equilibrium or it may be 99.9% or more complete.

Let us compare the hydrolysis of the two organic phosphates **adenosine triphosphate** (ATP) and **glucose 6-phosphate** (Eqs. 6-42 and 6-43).

$$HATP^{3-} + H_2O \longrightarrow HADP^{2-} + H_2PO_4^-$$
$$\Delta G° \approx -32.9\ kJ\ mol^{-1}$$
$$\Delta H° \approx -22.6\ kJ\ mol^{-1}\ \text{at}\ 25°C,\ \mu = 0.25 \qquad (6\text{-}42)$$

$$\alpha\text{-D-Glucose 6-phosphate}^- + H_2O \longrightarrow \text{glucose} + H_2PO_4^-$$
$$\Delta G° = -16.4\ kJ\ mol^{-1}\ \text{at}\ 25°C \qquad (6\text{-}43)$$

The decrease in Gibbs energy upon hydrolysis is twice as large for ATP as it is for glucose 6-phosphate. Glucose phosphate is thermodynamically more stable than ATP. It would be easier to form than would ATP by a reversal of the hydrolysis reaction and also easier to form biosynthetically. From the Gibbs energies of hydrolysis it follows that a **phospho group** could be transferred spontaneously from ATP to glucose in the presence of a suitable catalyst but not vice versa.



Phospho (phosphoryl) group

Because it reflects quantitatively the thermodynamic tendency for a group to be transferred to another

**TABLE 6-5**
**Values of p$K_a$, $\Delta G°$, and $\Delta H°$ for Ionization of Acids at 25°C [a,b]**

| Acid | p$K_a$ | $\Delta G°$ (kJ mol$^{-1}$) | $\Delta H°$ (kJ mol$^{-1}$) |
|---|---|---|---|
| Formic acid | 3.75 | 21.4 | 0.04 |
| Acetic acid | 4.76 | 27.2 | −0.1 |
| Propionic acid | 4.87 | 27.8 | −0.6 |
| Lactic acid | 3.97 (35°C) | 23.4 | 2.2 |
| Pyruvic acid | 2.49 | 14.2 | 12.1 |
| $NH_4^+$ | 9.25 | 52.8 | 52.2 |
| $CH_3NH_4^+$ | 10.59 | 60.4 | 55.4 |
| Alanine | | | |
| −COOH | 2.35 | 13.4 | 3.1 |
| $-NH_3^+$ | 9.83 | 56.1 | 45.4 |
| β-Alanine | | | |
| −COOH | 3.55 | 20.3 | 4.5 |
| $-NH_3^+$ (apparent) | 10.19 | 58.2 | |
| L-Alanyl-L-alanine | 3.34 | 19.1 | −0.5 |
| Aspartic acid | | | |
| −COOH | 2.05 | 11.7 | 7.7 |
| −COOH | 3.87 | 22.1 | 4.0 |
| $-NH_3^+$ | 10.60 | 60.5 | 38.8 |
| $H_2CO_3$, p$K_1$ | 6.35[c] | 36.2 | 9.4 |
| p$K_2$ | 10.33 | 59.0 | 15.1 |
| $H_3PO_4$, p$K_1$ | 2.12 | 12.1 | −7.9 |
| p$K_2$ | 7.18[d] | 41.0 | 3.8 |
| (apparent) | 6.78[e] | 38.7 | 3.3 |
| p$K_3$ | 12.40 | 70.8 | 17.6 |
| Glycerol 1-phosphate, p$K_2$ | 6.66 | 38.0 | −3.1 |
| Glucose 6-phosphate | 6.50 | 37.1 | −1.8 |
| Pyrophosphoric acid, $H_4P_2O_7$ | | | |
| p$K_3$ | 6.7 | 38.1 | −1.3 |
| (apparent) | 6.12[e] | 34.9 | 0.5 |
| p$K_4$ | 9.4 | 53.6 | −7.1 |
| (apparent) | 8.95[e] | 51.2 | 1.7 |
| Adenosine | 3.5 | 20.1 | 13.0 |
| AMP | | | |
| p$K_1$ (ring, apparent) | 3.74[e] | 21.3 | 4.2 |
| p$K_2$ (phosphate) | 6.67[d] | 38.1 | 3.6 |
| (apparent) | 6.45[e] | 36.8 | 3.6 |
| ADP | | | |
| p$K_2$ (ring, apparent) | 3.93[e] | 22.4 | 4.2 |
| p$K_3$ (diphosphate) | 7.20[d] | 41.1 | −5.7 |
| (apparent) | 6.83[f] | 39.0 | −5.7 |
| ATP | | | |
| p$K_3$ (ring, apparent) | 4.06 | 23.2 | 0 |
| p$K_4$ (triphosphate) | 7.68[d] | 43.8 | −7.0 |
| (apparent) | 7.06[d] | 40.2 | −7.0 |
| Pyridine | 5.17 | 29.5 | 20.1 |
| Phenol | 9.98 | 56.9 | 23.6 |

[a] These are thermodynamic values (infinite dilution) except for those labeled apparent. The latter apply at an ionic strength of 0.2–0.25.
[b] Most data are from Jencks, W. P. and Regenstein, J. (1976) in *Handbook of Biochemistry and Molecular Biology*, 3rd ed., Vol. I (Fasman, G. D. ed.), pp. 305–351, CRC Press, Cleveland, Ohio.
[c] Here, p$K_1$ is for $K_1 = [H^+][HCO_3^-] / [CO_2] + [H_2CO_3]$. From Forster, R. E., Edsall, J. T., Otis, A. B., and Roughton, F. J. W., eds. (1969) *NASA Spec. Publ.* **188**.
[d] From Phillips, R. C., George, P., and Rutman, R. J. (1963) *Biochemistry* **2**, 501–508.
[e] From Alberty, R. A. (1969) *J. Biol. Chem.* **244**, 3290–3302.
[f] Values used by Alberty, R. A. (1972) *Horizons of Bioenergetics*, pp. 135–147, Academic Press, New York, calculated for 0.2 ionic strength from equations of Phillips, R. C., George, P., and Rutman, R. J. (1966) *J. Am. Chem. Soc.* **88**, 2631–2640.

nucleophile (see Chapter 12), *the Gibbs energy decrease (−ΔG°) upon hydrolysis* is sometimes called the **group transfer potential**. During the hydrolysis of ATP (Eq. 6-42) the phospho group of ATP is transferred to a hydroxyl ion from water with a value $\Delta G° = -32.9$ kJ mol$^{-1}$. The group transfer potential of this phospho group is 32.9 kJ/mol and that of the phospho group of glucose 6-phosphate is only 27.6 kJ mol$^{-1}$. While the choice of water as the reference nucleophile for expression of the group transfer potential is somewhat arbitrary, it is customary. Transfer of groups is important in energy metabolism and in biosynthesis of polymers. Gibbs energies of hydrolysis are given in Table 6-6 for several compounds.

## 4. "Constants" That Vary with pH and Magnesium Ion Concentrations

Equation 6-42 is written for hydrolysis of $HATP^{3-}$ to $HADP^{2-} + H_2PO_4^-$, a stoichiometry that applies well in the pH range around 6. However, at a pH above ~7 most of the ATP is in the form $ATP^{4-}$ and is cleaved to $HPO_4^{2-}$ according to Eq. 6-44.

$$ATP^{4-} + H_2O \rightarrow ADP^{3-} + HPO_4^{2-} + H^+$$
$$\Delta G° \approx +5.4 \text{ kJ mol}^{-1} \text{ at } 25°C, \mu = 0.25$$
$$\Delta H° \approx -19.7 \text{ kJ mol}^{-1} \qquad (6\text{-}44)$$

The value of $\Delta G° = +5.4$ kJ mol$^{-1}$ for this reaction is hardly the large negative number expected for a highly spontaneous reaction. What is the matter? The problem is that $H^+$ is produced and that the standard state of $H^+$ is 1 M, not $10^{-7}$ M. Because of this, biochemists often prefer to use another kind of **apparent dissociation constant** and an **apparent ΔG** such that the standard state of $H^+$ is taken as that of the pH at which the experiments were done, usually pH 7. The symbol $K'$ has often be used and is used in this book to represent the following

pH-dependent equilibrium constant (Eq. 6-45) which will be a constant only at a single pH.

$$K' = \frac{[ADP^{3-}][HPO_4^{2-}]}{[ATP^{4-}]} \qquad (6\text{-}45)$$

If one proton is produced in the reaction as in Eq. 6-44, the following relationship will hold.

$$\Delta G' = \Delta G^\circ - 5.708 \times pH \text{ kJ mol}^{-1} \text{ at } 25^\circ C \qquad (6\text{-}46)$$

Note that $\Delta G' = -RT \ln K'$ and that $[H^+]$ does not appear in the expression for $K'$ given by Eq. 6-45. From the value $\Delta G^\circ = +5.4$ kJ mol$^{-1}$ and applying Eq.

**TABLE 6-6**
**Gibbs Energies of Hydrolysis at 25°C (in kJ mol$^{-1}$)[a]**

| Compound | Products | $\Delta G^\circ$ | $\Delta G'$(pH 7) | $\Delta H^\circ$ |
|---|---|---|---|---|
| $ATP^{4-}$ | $ADP^{3-} + HPO_4^{2-} + H^+$ | 5.41[b] | –34.54 | –19.71 |
| $ATP^{4-}$ | $AMP^{2-} + HP_2O_7^{3-} + H^+$ | 2.54[c] | –37.4 | –19.0 |
| $MgATP^{2-}$ | $MgADP^- + HPO_4^{2-} + H^+$ | 16.0[d] | –24.0 | –14.2 |
| $ADP^{3-}$ | $AMP^{2-} + HPO_4^{2-} + H^+$ | 3.67[c] | –36.3 | –13.5 |
| $AMP^{2-}$ | Adenosine + $HPO_4^{2-}$ | –9.6[e] | –9.6 | 0 |
| $ATP^{4-}$ | Adenosine + $HP_3O_{10}^{4-}$ | –36.0[e] | –36.0 | –7.9 |
| $HP_2O_7^{3-}$ | $2\,HPO_4^{2-} + H^+$ | 6.54[c] | –33.4 | –12.6 |
| Acetyl phosphate$^{2-}$ | Acetate$^-$ + $HPO_4^{2-} + H^+$ | –7.7[f] | –47.7 | |
| 1,3-Diphosphoglycerate$^{4-}$ | 3-Phosphoglycerate$^{3-}$ + $HPO_4^{2-} + H^+$ | –14.5[g] | –54.5 | |
| Phosphoenolpyruvate$^{3-}$ | Pyruvate$^-$ + $HPO_4^{2-}$ | –61.9 | –61.9 | –25.1 |
| Carbamoyl phosphate$^{2-}$ | $H_2N-\overset{\overset{O}{\|}}{C}-O^- + HPO_4^{2-} + H^+$ | –11.5 | –51.5 | |
| Creatine phosphate$^-$ | Creatine$^+$ + $HPO_4^{2-}$ | –43.1 | –43.1 | |
| Phosphoarginine$^-$ | Arginine + $HPO_4^{2-}$ | | –38.1[h] | ($Mg^{2+}$ present) |
| Glycerol phosphate$^{2-}$ | Glycerol + $HPO_4^{2-}$ | –9.2 | –9.2 | |
| α-D-Glucose 6-phosphate$^{2-}$ | α-D-Glucose + $HPO_4^{2-}$ | –13.8 | –13.8 | –2.5 |
| Glucose 1-phosphate$^{2-}$ | Glucose + $HPO_4^{2-}$ | –20.9 | –20.9 | |
| Maltose (or glycogen) | 2-Glucose | –16.7 | –16.7 | |
| Sucrose | Glucose + fructose | –29.3 | –29.3 | |
| UDP glucose$^{2-}$ | Glucose + $UDP^{3-} + H^+$ | 9.4 | –30.5 | |
| $N^{10}$-Formyltetrahydrofolic acid | Formate$^-$ + $H^+$ + tetrahydrofolic acid | 14.1 | –25.9[h] | |
| Acetic anhydride | 2-Acetate$^-$ + $2H^+$ | 31.1 | –48.9 | |
| Acetyl-CoA | Acetate$^-$ + $H^+$ + CoA | 4.9 | –35.1[i] | |
| Succinyl-CoA$^-$ | Succinate$^{2-}$ + $H^+$ + CoA | –3.5 | –43.5[j] | |
| Ethyl acetate | Ethanol + acetate$^-$ + $H^+$ | 20.2 | –19.7 | |
| Asparagine | Aspartate$^-$ + $NH_4^+$ | –15.1 | –15.1 | |
| Glycine ethyl ester$^+$ (39°C) | Glycine + ethanol + $H^+$ | 4.9 | –35.1 | |
| Valyl-tRNA$^+$ | Valine + tRNA + $H^+$ | 4.9 | –35.1 | |

[a] Unless indicated otherwise, the values are based on tables from Jencks, W. P. (1976) *Handbook of Biochemistry and Molecular Biology*, 3rd ed., Vol I (Fasman,G.D. ed.), pp. 296–304. CRC Press, Cleveland, Ohio. For a reaction producing one proton at pH 7 $\Delta G' = -39.96$ kJ mol$^{-1.}$

[b] Guynn, R.W. and Veech, R.L. (1973) *J. Biol. Chem.* **248**, 6966–6972.

[c] Based on +11.80 kcal mol$^{-1}$ for hydrolysis to $P_2O_7^{4-}$ plus $\Delta G^\circ$ of dissociation of $HP_2O_7^{3-}$ as quoted by Alberty, R. A. (1969) *J.Biol.Chem.* **244**, 3290–3324. However, 1.017 kcal mol$^{-1}$ was added to the value of 11.80 to make it consistent with that for hydrolysis of ATP to ADP. Reevaluation by Frey and Arabshahi (1995) *Biochemistry* **34**, 11307–11310, indicates that $\Delta G'$ (pH 7) for hydrolysis of ATP to AMP + $PP_i$ is ~ 10 kJ mol$^{-1}$ more negative than is shown here.

[d] Alberty, R.A. (1972) *Horizons of Bioenergetics*, Academic Press, New York, pp. 135–147.

[e] George, P. , Witonsky, R.J., Trachtman, M., Wu, C., Dorwart, W., Richman, L., Richman, W., Shurayh, F., and Lentz, B. (1970) *Biochim. Biophys. Acta.* **223**, 1–15.

[f] Based on $\Delta G^+ = 3.0$ kcal mol$^{-1}$ for acetyl phosphate + CoA → acetyl-CoA + $P_i$ from Stadtman, E. R. (1973) *The Enzymes*, (Boyer, P.D., ed.), 3rd ed., Vol. 8. pp. 1–49. Academic Press, New York, together with $\Delta G'$ (pH 7) for hydrolysis of acetyl-CoA.

[g] Estimated from $\Delta G^\circ$ for ATP hydrolysis + $\Delta G' = -19.9$ kJ mol$^{-1}$ for the 3-phosphoglycerate kinase reaction: Burton, K. and Krebs, H. A. (1953) *Biophys. J.* **54**, 94–107; and (1955) *Biophys. J.* **59**, 44–46.

[h] Estimated from data at pH 7.7 or 8.0 (tables of Jencks).

[i] Guynn, R. W., Gelberg, H. J., and Veech, R. L. (1973) *J. Biol. Chem.* **248**, 6957–6965 found $\Delta G^\circ = -35.75$ kJ mol$^{-1}$ at 38°C. Without data on $\Delta H$, correction to 25°C is difficult. Burton, K., (1955) *Biophys. J.* **59**, 44–46, gave $\Delta G'$ (pH 7) for $ATP^{4-}$ + acetate$^-$ + CoA → $ADP^{3-}$ + acetyl-CoA + $HPO_4^{2-}$ as approximately zero at 25°C. Guynn *et al.* found –0.56 kJ mol$^{-1}$ at 38°C. This same value (–0.56 kJ) at 25°C was assumed to obtain the figure given here. This is equivalent to assuming $\Delta H^\circ$ of hydrolysis as almost the same for ATP and acetyl-CoA, an unsupported assumption.

[j] Assumed 2 kcal mol$^{-1}$ more negative than that of acetyl-CoA as in tables of Jencks.

6-46, we obtain for the hydrolysis of ATP at 25°C, μ = 0.25:

$$\Delta G'\ (\text{pH } 7) = -34.5\ \text{kJ mol}^{-1} = -8.26\ \text{kcal mol}^{-1} \qquad (6\text{-}47)$$

An additional set of standard states is frequently met in the biochemical literature. An equilibrium constant, designated in this book as $K^{\dagger}$, is used to relate the *total concentrations of all ionic forms* of the components present at the pH of the experiment. Thus,

$$K^{\dagger} = \frac{[\text{ADP, all forms}][\text{phosphate, all forms}]}{[\text{ATP, all forms}]} \qquad (6\text{-}48)$$

## BOX 6-A MEASUREMENT OF INTRACELLULAR pH

What is the pH within a cell? Is it constant or does it vary with physiological conditions? Do all cells operate at similar pH? The answers to these important questions have been sought using a variety of techniques.[a] Tiny microelectrodes with tips only 1 μm in diameter have been inserted into cells. Indicator dyes have been diffused into cells and either light absorption or fluorescence[b] measured. The distribution of a suitable radiolabeled weak acid or weak base, such as [$^{14}$C]methylamine, that is able to permeate cells can be used to calculate the difference between internal and external pH.[c] The activity of the enzyme carbonic anhydrase, which is very pH sensitive, can be used to monitor the pH of mitochrondia.[d] Since 1973 NMR methods have been popular.[e–l] The chemical shifts of $^{31}$P in inorganic phosphate (p$K_a$ = 6.9), ATP, glucose 6-phosphate, and of other metabolites of $^{13}$C in citrate and bicarbonate,[l] give direct estimates of pH. However, caution must be exercised if internal pH values good to ± 0.1 unit are to be obtained.[g] More sensitive measurements over the pH range 1.3 – 9.1 can be made by diffusing one of a series of aminophosphonates into cells and measuring the $^{31}$P chemical shift.[h] Fluorinated probes such as dimethylfluoroalanine (p$K_a$ = 7.3) are also useful because of their low toxicity and high sensitivity of the $^{19}$F NMR signal. These alanine derivatives can be diffused into cells as their methyl esters, which are rapidly cleaved within cells, allowing the fluorinated amino acids to accumulate.[i,m]

The pH within cells appears to be tightly controlled although small variations are sometimes observed. Red blood cells, thymocytes, liver, skeletal muscles, and intact hearts all maintain a pH in the range 7.0 – 7.3.[b,h,i] However, the pH can fall to 6.2 within 13 minutes of oxygen deprivation (ischemia) and to 6.1 after exhaustive exercise.[n,o] The $^{31}$P NMR technique permits the monitoring of pH as well as the state of the adenylate system (Section D) in human limbs suffering from circulatory insufficiency.[c]

In higher plants the pH of cytoplasm is 7.4–7.5 but vacuoles are acidic with a pH of 4.5–6.[l] The cytoplasm of maize root tips has a pH of 7.1 but the vacuoles are at a pH of 5.5.[f] The pH of granules of the chromaffin cells of the adrenal cortex, which accumulate high concentrations of ATP and catecholamines, is also low, ~5.7.[p] The bacterium *Streptococcus faecalis* maintains a higher internal pH of ~8.0, even when the pH of the medium varies from 6.5 to 8.0[q] while *E. coli* operates at pH 7.6, the extremes of variation being 7.4 – 7.8 when the external pH changes from 5.5–9.[r]

Changes of internal pH during developmental events such as fertilization of sea urchin eggs (+0.3 unit) have been recorded. However, the significance of pH changes in metabolic regulation remains uncertain.[c]


[a] Kotyk, A., and Slavik, J. (1989) *Intracellular pH and its Measurement*, CRC Press, Boca Raton, Florida
[b] Rogers, J., Hesketh, T. R., Smith, G. A., and Metcalfe, J. C. (1983) *J. Biol. Chem.* **258**, 5994–5997
[c] Nuccitelli, R., and Deamer, D. W., eds. (1982) *Intracellular pH: Its Measurement, Regulation and Utilization in Cellular Functions*, Liss, New York
[d] Dodgson, S. J., Forster, R. E., II, and Storey, B. T. (1982) *J. Biol. Chem.* **257**, 1705–1711
[e] Moon, R. B., and Richards, J. H. (1973) *J. Biol. Chem.* **248**, 7276–7278
[f] Roberts, J. K. M., Ray, P. M., Wade-Jardetsky, N., and Jardetsky, O. (1980) *Nature (London)* **283**, 870–872
[g] Avison, M. J., Hetherington, H. P., and Shulman, R. G. (1986) *Ann. Rev. Biophys. Biophys. Chem.* **15**, 377–402
[h] Pietri, S., Miollan, M., Martel, S., Le Moigne, F., Blaive, B., and Culcasi, M. (2000) *J. Biol. Chem.* **275**, 19505–19512
[i] Taylor, J. S., and Deutsch, C. (1983) *Biophys. J.* **43**, 261–267
[j] Bailey, I. A., Williams, S. R., Radda, G. K., and Gadian, D. G. (1981) *Biochem. J.* **196**, 171–178
[k] Barton, J. K., Den Hollander, J. A., Lee, T. M., MacLaughlin, A., and Shulman, R. G. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2470–2473
[l] Gout, E., Bligny, R., and Douce, R. (1992) *J. Biol. Chem.* **267**, 13903–13909
[m] Deutsch, C. J., and Taylor, J. S. (1987) *Ann. N.Y. Acad. Sci.* **508**, 33
[n] Garlick, P. B., Radda, G. K., and Seeley, P. J. (1979) *Biochem. J.* **184**, 547–554
[o] Pan, J. W., Hamm, J. R., Rothman, D. L., and Shulman, R. G. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7836–7839
[p] Pollard, H. B., Shindo, H., Creutz, C. E., Pazoles, C. J., and Cohen, J. S. (1979) *J. Biol. Chem.* **254**, 1170–1177
[q] Kobayashi, H., Murakami, N., and Unemoto, T. (1982) *J. Biol. Chem.* **257**, 13246–13252
[r] Slonczewski, J. L., Rosen, B. P., Alger, J. R., and MacNab, R. M. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6271–6275

and

$$\Delta G^{\dagger} = -RT \ln K^{\dagger} \qquad (6\text{-}49)$$

The Gibbs energy change $\Delta G^{\dagger}$ can be related to $\Delta G'$ by considering the relationship of $K^{\dagger}$ to $K'$. For ATP hydrolysis in the pH range of 2–10, $K^{\dagger}$ is given by Eq. 6-50.

$$K^{\dagger} = \frac{K' \left(1 + \frac{[H^+]}{K_{HADP^{2-}}} + \frac{[H^+]^2}{K_{HADP^{2-}} K_{H_2ADP^-}}\right)\left(1 + \frac{[H^+]}{K_{HPO_4^{2-}}}\right)}{\left(1 + \frac{[H^+]}{K_{HATP^{3-}}} + \frac{[H^+]^2}{K_{HATP^{3-}} K_{H_2ATP^{2-}}}\right)} \qquad (6\text{-}50)$$

In this equation $K_{HADP^{2-}}$, etc., are consecutive dissociation constants as given in Table 6-4. The expressions in parentheses are the **Michaelis pH functions**, which were considered in Chapter 3 (Eqs. 3-4 to 3-6). In Eq. 6-50 they relate the total concentration of each component to the concentration of the most highly dissociated form. Thus, for the pH range 2–10

$$[P_i]_{total} = [HPO_4^{2-}](1 + [H^+] / K_{H_2PO_4^-}) \qquad (6\text{-}51)$$

Using apparent $pK_a$ values ($\mu = 0.2$) for $H_2PO_4^-$, $HADP^{2-}$, and $HATP^{3-}$ of 6.78, 6.83, and 7.06 (Table 6-5) and taking $\Delta G'$ at pH 7 as –34.5 kJ mol$^{-1}$, we compute $\Delta G^{\dagger} = -35.0$ kJ mol$^{-1}$ at pH 7. The difference between $\Delta G'$ and $\Delta G^{\dagger}$ in this case is small, but it would be larger if the ionic forms in Eq. 6-44 were not the ones predominating at pH 7.

To obtain the Gibbs energy change for a reaction under *other than standard conditions,* Eq. 6-29 must be applied. Thus, at pH 7 and 0.01 M activities of $ADP^{3-}$, $ATP^{4-}$, and $HPO_4^{2-}$, $\Delta G$ for hydrolysis of ATP according to Eq. 6-44 is $-34.5 - (2 \times 5.71) = -45.9$ kJ mol$^{-1}$ = –11.0 kcal mol$^{-1}$. We see that at concentrations existing in cells (usually in the millimolar range) ATP has a substantially higher group transfer potential than under standard conditions.

The reader should bear in mind that there is no accepted standard usage of $K'$ and that $K^{\dagger}$ is just for this book! An international committee[21] has recommended that $K'$ be used with the same meaning as $K^{\dagger}$ in this book and more changes may be coming (see the next section).

To obtain $\Delta G$ at a temperature other than 25°C, we must know $\Delta H$ for the reaction. Using Eq. 6-37a it is easy to show that $\Delta G$ at temperature $T_2$ is related to that at $T_1$ as follows:

$$\Delta G_2 \approx \frac{T_2 \, \Delta G_1 - (T_2 - T_1) \, \Delta H}{T_1} \qquad (6\text{-}52)$$

The enthalpy of hydrolysis of ATP according to Eq. 6-44 is approximately –19.7 kJ mol$^{-1}$. Using this value and applying Eq. 6-52 we can calculate that $\Delta G'$ (pH 7) for the hydrolysis of ATP at 38°C is –35.2 kJ mol$^{-1}$. The value of $\Delta G'$ (25°C, pH 7), used in obtaining this answer, is – 35.54 kJ mol$^{-1}$ which was computed from the value of $\Delta G'$ (pH 7), 38°C of 35.19 kJ mol$^{-1}$ reported by Guynn and Veech[37] using Eq. 6-52. Bear in mind that all of the foregoing Gibbs energy changes are *apparent* values applying to solutions of ionic strength ~0.25.

Both ADP and ATP as well as inorganic pyrophosphate form complexes with metal ions. Since the magnesium complexes are often the predominant forms of ADP and ATP under physiological conditions, we must consider the following Gibbs energy changes. These are apparent values for $\mu = 0.2$ at 25°C.

$$ATP^{4-} + Mg^{2+} \rightarrow Mg\,ATP^{2-}$$
$$\Delta G^{\circ} = -26.3; \; \Delta H^{\circ} = 13.8 \text{ kJ mol}^{-1} \qquad (6\text{-}53)$$

$$ADP^{3-} + Mg^{2+} \rightarrow Mg\,ADP^{-}$$
$$\Delta G^{\circ} = -19.8; \; \Delta H^{\circ} = 15.1 \text{ kJ mol}^{-1} \qquad (6\text{-}54)$$

Combining the apparent $\Delta G^{\circ}$ values for Eqs. 6-47, 6-53, and 6-54, we obtain

$$MgATP^{2-} + H_2O \rightarrow MgADP^{-} + HPO_2^{2-} + H^+$$
$$\Delta G' \text{ (pH 7)} = -28.0 \text{ kJ mol}^{-1} \qquad (6\text{-}55)$$

The stoichiometry of Eq. 6-55 never holds exactly. Some of the $Mg^{2+}$ dissociates from the $MgADP^-$; both protons and $Mg^{2+}$ bind to $H_2PO_4^-$; and $HATP^{3-}$ and $HADP^{2-}$ are present and bind $Mg^{2+}$ weakly.[38] Thus, the observed value of $\Delta G^{\dagger}$ varies with both pH and magnesium concentration as well as with changes in ionic strength. Tables and graphs showing the apparent value of $\Delta G^{\dagger}$ under various conditions have been prepared by Alberty[38] and by Phillips *et al.*[39] An example is shown in Fig. 6-2. From this graph we find that $\Delta G^{\dagger}$ for hydrolysis of ATP at pH 7, 25°C, $\mu = 0.2$, and 1 mM $Mg^{2+}$ (a relatively high intracellular concentration[40]) is –30.35 kJ mol$^{-1}$ (–7.25 kcal mol$^{-1}$).

Figure 6-2 was drawn using the equations of Alberty, but the value of $\Delta G'$ (pH 7) of hydrolysis of ATP = 34.54 kJ mol$^{-1}$ at $[Mg^{2+}] = 0$ based on results of Guynn and Veech[37] was used. The $pK_a$ values and formation constants of $Mg^{2+}$ complexes were those of Alberty.[38] Note that George *et al.*[40] provided formation constants of these complexes at infinite dilution where the values of $\Delta G^{\circ}$ of formation are considerably more negative than those given in Eqs. 6-53 to 6-55.

From the foregoing considerations we see that complexing with $Mg^{2+}$ somewhat decreases the group transfer potential of the phospho group of ATP. Furthermore, changes in the concentration of free $Mg^{2+}$ with time and between different regions of a cell may

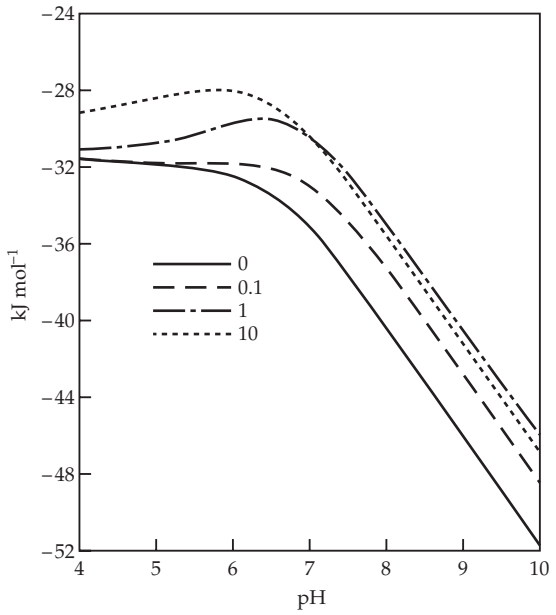



**Figure 6-2** Plots of the apparent Gibbs energy $\Delta G^{\dagger}$ for hydrolysis of ATP as a function of pH at a series of different concentrations of free magnesium ions. Millimolar $[Mg^{2+}]$ is indicated by the numbers by the curves. Computer-drawn graphs courtesy of Carol M. Metzler.

have significant effects.[41] While $Mg^{2+}$ is a principal cation in tissues, it is by no means the only one. Thus, $Ca^{2+}$, $Mn^{2+}$, and even $K^{+}$ will affect equilibria involving polyphosphates such as ATP.

It is not easy to measure the group transfer potential of ATP and published values vary greatly. George *et al.*[40] reported $\Delta G°$ for Eq. 6-42 as −39.9 kJ mol$^{-1}$ at 25°C, μ = 0.2, and as − 41.25 kJ mol$^{-1}$ at infinite dilution. They regarded these values as good to ±4 kJ mol$^{-1}$. However, most other estimates have been at least 4 kJ mol$^{-1}$ less negative.[42–44] The self-consistent set of thermodynamic data used throughout this book are based in part on the value of ΔG for hydrolysis of ATP obtained by Guynn and Veech.[37]

## 5. A New Standard for Biochemical Thermodynamics?

Because of the complexities of the equilibria involved in biochemical reactions Alberty and others[45–47] and a Panel on Biochemical Thermodynamics[48] have suggested that tables of thermodynamic properties for biochemical use be tabulated for the following conditions: $T$ = 298.15 K, $P$ = 1 bar (0.1 $MP_a \equiv 0.987$ atm), pH = 7, pMg = 3 ($[Mg^{2+}] = 10^{-3}$ M), and ionic strength I (μ in this chapter) = 0.25 M. There appears to be both advantages and disadvantages to this. The proponents suggest that the symbols $\Delta G'$, $K'$, etc. be used for this new standard. To change all of the numbers in this book, which uses a self-consistent set of thermodynamic quantities, is impractical. However, it is worthwhile to compare the value of Alberty and Goldberg for $\Delta G'°$ for hydrolysis of ATP to ADP + $P_i$ under the proposed biochemical standard conditions with other values given in this chapter, all at 25°C.

$\Delta G'°$ = −32.49 kJ mol$^{-1}$ (new proposed biochemical standard)
$\Delta G^{\dagger}$ (pH 7) −35.0 kJ mol$^{-1}$ (Eq. 6-50)
$\Delta G'$ (pH 7) −34.5 kJ mol$^{-1}$ (for $ATP^{4-} \rightarrow ADP^{3-} + HPO4^{2-}$; Eq. 6-47)
$\Delta G'$ (pH 7) −28.0 kJ mol$^{-1}$ (for $MgATP^{2-} \rightarrow MgADP^{-} + HPO4^{2-}$; Eq. 6-55)

## 6. Bond Energies and Approximate Methods for Estimation of Thermodynamic Data

For approximate estimation of enthalpy changes during reactions, use can be made of empirical bond energies (Table 6-7) which represent the approximate enthalpy changes ($-\Delta H°$) for formation of compounds in a gaseous state from atoms in the gas phase. Other more comprehensive methods of approximation have been developed.[49,50]

## 7. Gibbs Energies of Combustion by $O_2$ and by $NAD^+$

Since oxidation processes are so important in the metabolism of aerobic organisms, it is often convenient to discuss Gibbs energies of combustion. These are easily derived from the Gibbs energies of formation. For example, $\Delta G_c$ for acetate ion (aqueous) may be obtained as follows:

$$Acetate^- + H^+ \rightarrow 2\,C + 2\,H_2 + O_2$$
$\Delta G° = +369.2$ kJ mol$^{-1}$
$= -\Delta G_f°$ of acetate (from Table 6-3)

$$2\,H_2 + O_2 \rightarrow 2\,H_2O$$
$\Delta G° = 2 \times -237.2 = -474.4$ kJ mol$^{-1}$
$= 2 \times \Delta G_f°$ of $H_2O$ (6-56)

$$2\,C + 2\,O_2 \rightarrow 2\,CO_2$$
$\Delta G° = 2 \times -394.4 = -788.8$ kJ mol$^{-1}$
$= 2 \times \Delta G_f°$ of $CO_2$

$$Acetate^-\ (1\ M) + H^+\ (1\ M) + 2\,O_2\ (1\ atm) \rightarrow 2\,CO_2\ (1\ atm) + H_2O\ (l)$$
$\Delta G° = -894.0$ kJ mol$^{-1}$
$= \Delta G_c°$ of $acetate^-$

**TABLE 6-7**
**Empirical Bond Energies and Resonance Energies[a]**

| Energy values for single bonds (kJ mol$^{-1}$) | | Energy values for multiple bonds (kJ mol$^{-1}$) | | Empirical resonance energy values (kJ mol$^{-1}$) | |
|---|---|---|---|---|---|
| H–H | 436 | C=C | 615 | Benzene | 155 |
| C–C | 346 | N=N | 418 | Naphthalene | 314 |
| Si–Si | 177 | O=O | 402 ('Δ state) | | |
| N–N | 161 | C=N | 615 | Styrene | 155 + 21 |
| O–O | 139 | C=O | 686 (formaldehyde) | Phenol | 155 + 29 |
| S–S | 213 | | 715 (aldehydes) | | |
| C–H | 413 | | 728 (ketones) | Benzaldehyde | 155 + 17 |
| Si–H | 295 | C=S | 477 | Pyridine | 180 |
| N–H | 391 | C≡C | 812 | Pyrrole | 130 |
| P–H | 320 | N≡N | 946 ($N_2$) | Indole | 226 |
| O–H | 463 | C≡N | 866 (HCN) | | |
| S–H | 339 | | 891 (nitriles) | —C(=O)OH | 117 |
| C–Si | 290 | | | | |
| C–N | 292 | | | —C(=O)OR | 100 |
| C–O | 351 | | | | |
| C–S | 259 | | | —C(=O)$NH_2$ | 88 |
| C–F | 441 | | | | |
| C–Cl | 328 | | | Urea | 155 |
| C–Br | 276 | | | CO | 439 |
| C–I | 240 | | | $CO_2$ | 151 |
| Si–O | 369 | | | | |

[a] From Pauling, L. ( 1960) *The Nature of the Chemical Bond*, 3rd ed., pp. 85, 189, and 195–198, Cornell Univ. Press, Ithaca, New York

Table 6-4 lists values of $\Delta G_c^\circ$ as well as those of $\Delta G_f^\circ$. Besides $CO_2$ and $H_2O$, the other products assumed in calculating the values of $\Delta G_c^\circ$ in Table 6-4 are $N_2$ (1atm), $H_3PO_4$ (1 M), and $H_2SO_4$ (1 M).

Much of the oxidation occurring in cells is carried out by the biological oxidizing agent **nicotinamide adenine dinucleotide** (**NAD$^+$**) or by the closely related $NADP^+$ (Chapter 10). It is convenient to tabulate values of $\Delta G^\circ$ for complete oxidation of compounds to $CO_2$ using $NAD^+$ rather than $O_2$ as the oxidizing agent. These values, designated $\Delta G^\circ{}'_{ox}$ and $\Delta G'_{ox}$ (pH 7), are also given in Table 6-3. Notice that these values are relatively small, corresponding to the fact that little energy is made available to cells by oxidation with $NAD^+$; for example (Eq. 6-57):

$$\text{Acetate}^- + 2\,H_2O + 4\,NAD^+ \rightarrow 2\,CO_2 + 4\,NADH + 3\,H^+\,(10^{-7}\,M)$$
$$\Delta G'_{ox}\,(\text{pH } 7) = +22.1 \text{ kJ mol}^{-1} \qquad (6\text{-}57)$$

When the reduced $NAD^+$ (designated NADH) so formed is reoxidized in mitochondria (Eq. 6-58), a large amount of energy is made available to cells.

$$4\,NADH + 4\,H^+ + 2\,O_2 \rightarrow 4\,H_2O + 4\,NAD^+$$
$$\Delta G'_{ox}\,(\text{pH } 7) = -876.1 \text{ kJ mol}^{-1} \qquad (6\text{-}58)$$

The sum of Eqs. 6-57 and 6-58 is the equation for combustion of acetate by $O_2$ (Eq. 6-59) and the two $\Delta G$ values sum to $\Delta G_c$ for acetate$^-$.

$$\text{Acetate}^- + H^+ + 2\,O_2 \rightarrow 2\,CO_2 + 2\,H_2O$$
$$\Delta G_c^\circ = -894.0 \text{ kJ mol}^{-1}$$
$$\Delta G_c'\,(\text{pH } 7) = -854.0 \text{ kJ mol}^{-1} \qquad (6\text{-}59)$$

The values of $\Delta G_{ox}$ (Table 6-4) not only give an immediate indication of the relative amounts of energy available from oxidation of substrate with $NAD^+$ but also are very convenient in evaluating $\Delta G$ for fermentation reactions. For example, consider the fermentation of glucose to ethanol (Eq. 6-60):

$$\alpha\text{-D-Glucose} \rightarrow 2\,CO_2 + 2\,\text{ethanol} \qquad (6\text{-}60)$$

The Gibbs energy change $\Delta G'$ (pH 7) for fermentation of glucose to ethanol and $CO_2$ can be written immediately from the data of Table 6-4 (Eq. 6-61).

$$\Delta G'\,(\text{pH } 7) = -243.8 - 2(-4.6)$$
$$= -234.6 \text{ kJ mol}^{-1} \qquad (6\text{-}61)$$

The values of $\Delta G_{ox}$ for $H_2O$, $CO_2$, and $H^+$ are always zero and need not be considered. The same computation can be made using the awkwardly large values of $\Delta G_c^\circ$ or using values of $\Delta G_f^\circ$ The latter are also large, and $CO_2$ and water must be considered in the equations. Table 6-4 can be used to obtain values of $\Delta G^\circ$ for many metabolic reactions considered later in the book. Data from any column in the table may be used for this purpose, but for simplicity try the values in the $\Delta G'_{ox}$ column.

Note, however, that for reactions involving oxidation by $O_2$ or by any oxidant, other than $NAD^+$, not appearing in Table 6-4, the following two-step procedure is necessary if the $\Delta G_{ox}$ values are used. From the $\Delta G_{ox}$ values compute $\Delta G^\circ$ or $\Delta G'$ for the reaction under consideration using $NAD^+$ as the oxidant. Then add to this the Gibbs energy of oxidation by $O_2$ (or other oxidant) of the appropriate number of moles of NADH formed. The latter is given for $O_2$ in Table 6-8 and can also be evaluated for a number of other oxidants, such as $Fe^{3+}$ and cytochrome *c*, from the data in Table 6-8.

## BOX 6-B MAGNESIUM

After potassium ion $Mg^{2+}$ is the most abundant cation present in tissues[a,b]. The average adult ingests ~10–12 mmol of magnesium ion daily (~1/4 g). Of this, about one-third is absorbed from the digestive tract. An equivalent amount is excreted in the urine to maintain homeostasis. Sixty percent of the magnesium in the body is found in the bones. The total $Mg^{2+}$ concentration in serum is ~1 mM, while various tissues contain as much as 5–17 mM.[c] The concentration of *free* $Mg^{2+}$ is difficult to measure. A variety of measurements using $^{31}P$ NMR spectroscopy of ATP,[d–f] $^{19}F$ NMR of added fluorine-containing chelators,[g] fluorescent chelators,[h] ion selective electrodes,[c] and other indirect procedures[g,i,j] indicate that the free magnesium concentration is almost the same in extracellular fluids, cytosol, and mitochondria.[h] Most values have been about 0.5 mM, somewhat less in erythrocytes. However, the most recent estimates indicate an intracellular [$Mg^{2+}$] of 0.8–1.1 mM.[c,f]

The additional $Mg^{2+}$ present in cells is bound to proteins, nuclei acids, and soluble compounds such as ATP, ADP, citrate,[k] and other phosphate- and carboxylate-containing substances. The binding is reversible and equilibration is usually rapid. It has been suggested that [$Mg^{2+}$], like [$H^+$], remains relatively constant within cells and that these two ions are in free equilibrium in the blood serum.[l] Nevertheless, there are instances in which at least temporary alterations in concentrations of both free $Mg^{2+}$ and $H^+$ occur.[m] During rapid catabolism of carbohydrates the formation of lactic acid by glycolysis leads to acidification of muscle cells, the pH falling from 7.3 to as low as 6.3. This drop in pH causes a large decrease in the extent of binding of $Mg^{2+}$ to molecules such as ATP and to a transient increase in [$Mg^{2+}$]. Similarly, the release of bisphosphoglycerate from hemoglobin upon oxygenation leads to a decreased concentration of free $Mg^{2+}$ as the latter coordinates with bisphosphoglycerate.[n] The 0.25 mM [$Mg^{2+}$] of aerobic red blood cells rises to 0.67 mM under anaerobic conditions.[d] Such changes in the free $Mg^{2+}$ concentrations will affect many equilibrium[o] and may be of significance in metabolic regulation.

The magnesium ion has a smaller radius than $Ca^{2+}$, a fact that may account for its more ready entry into cells. $Mg^{2+}$ can often be replaced by $Mn^{2+}$ with full activity for enzymes that require it. On the other hand, high concentrations of $Ca^{2+}$ are often antagonistic to $Mg^{2+}$. This antagonism is clearly seen in the effect of the two ions on irritability of protozoa.[p] Both deficiency of magnesium and excess of calcium in the surrounding medium cause increased irritability. Excess magnesium leads to anesthesia. The $Mg^{2+}$ concentration is high in hibernating animals.

Over 300 enzymes are dependent upon $Mg^{2+}$, the largest single group being the phosphotransferases, for which MgATP complex may be regarded as the substrate.[q] Magnesium has a special role in photosynthesis as a component of chlorophyll.

One of the most toxic metals is beryllium. It has been suggested that $Be^{2+}$ competes with $Mg^{2+}$ at many enzyme sites, including those of phosphoglucomutase and of phosphatases. However, as pointed out by Petsko,[r] because of its small size beryllium tends not to form $Be^{2+}$ ions but, rather, covalent complexes such as $BeF_3{\bullet}OH_2^-$, $BeF_4^{2-}$, and MgADP•-$BF_2$(OH). In the latter complex beryllium is covalently linked to the oxygen atom of the β phospho group of ADP to give an analog of MgATP, which inhibits many enzymes. See for example, Fig. 19-16A, in which MgADP•$BeF_x$ occupies the active site of myosin.

---


a Cowan, J. A., ed. (1995) *The Biological Chemistry of Magnesium*, VCH Publ., New York

b Strata, P., and Carbone, E., eds. (1991) *$Mg^{2+}$ and Excitable Membranes*, Springer-Verlag, Berlin and New York

c McGuigan, J. A. S., Blatter, L. A., and Buri, A. (1991) in *$Mg^{2+}$ and Excitable Membranes* (Strata, P., and Carbone, E., eds), Springer-Verlag, Berlin and New York

d Gupta, R. K., and Moore, R. D. (1980) *J. Biol. Chem.* **255**, 3987–3993

e Garfinkel, L., and Garfinkel, D. (1984) *Biochemistry* **23**, 3547–3552

f Clarke, K., Kashiwaya, Y., King, M. T., Gates, D., Keon, C. A., Cross, H. R., Radda, G. K., and Veech, R. L. (1996) *J. Biol. Chem.* **271**, 21142–21150

g Levy, L. A., Murphy, E., Raju, B., and London, R. E. (1988) *Biochemistry* **27**, 4041–4048

h Jung, D. W., Apel, L., and Brierley, G. P. (1990) *Biochemistry* **29**, 4121–4128

i Corkey, B. E., Duszynski, J., Rich, T. L., Matschinsky, B., and Williamson, J. R. (1986) *J. Biol. Chem.* **261**, 2567–2574

j Magneson, G. R., Puvathingal, J. M., and Ray, W. J., Jr. (1987) *J. Biol. Chem.* **262**, 11140–11148

k Kwack, H., and Veech, R. L. (1992) *Curr. Top. Cell. Regul.* **33**, 185–207

l Veloso, D., Guynn, R. W., Oskarrson, M., and Veech, R. L. (1973) *J. Biol. Chem.* **248**, 4811–4819

m Purich, D. L., and Fromm, H. J. (1972) *Curr. Top. Cell. Regul.* **6**, 131–167

n Bunn, H. F., Ransil, B. J., and Chao, A. (1971) *J. Biol. Chem.* **246**, 5273–5279

o Cornell, N. W. (1979) *J. Biol. Chem.* **254**, 6522–6527

p Meli, J., and Bygrave, F. L. (1972) *Biochem. J.* **128**, 415–420

q Vink, R., McIntosh, T. K., and Faden, A. I. (1991) in *$Mg^{2+}$ and Excitable Membranes* (Strata, P., and Carbone, E., eds), Springer-Verlag, Berlin and New York

r Petsko, G. A. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 538–540

## C. Electrode Potentials and Gibbs Energy Changes for Oxidation–Reduction Reactions

We live under a blanket of the powerful oxidant $O_2$. By cell respiration oxygen is reduced to $H_2O$, which is a very poor reductant. Toward the other end of the scale of oxidizing strength lies the very weak oxidant $H^+$, which some bacteria are able to convert to the strong reductant $H_2$. The $O_2$–$H_2O$ and $H^+$ – $H_2$ couples define two biologically important oxidation–reduction (**redox**) systems. Lying between these two systems are a host of other pairs of metabolically important substances engaged in oxidation–reduction reactions within cells.

There are two common methods for expressing the oxidizing or reducing powers of redox couples in a quantitative way. On the one hand, we can list values of $\Delta G$ for oxidation of the reduced form of a couple to the oxidized form by $O_2$. A compound with a large value of $-\Delta G$ for this oxidation will be a good reductant. An example is $H_2$ for which $\Delta G$ of combustion at pH 7 (Table 6-4) is – 237 kJ / mol. Poor reductants such as $Fe^{2+}$ are characterized by small values of $\Delta G$ of oxidation ($-8.5$ kJ mol$^{-1}$ for $2\ Fe^{2+} \rightarrow 2\ Fe^{3+}$). The Gibbs energies of oxidation of biological hydrogen carriers, discussed in Chapter 15, for the most part fall between those of $H_2$ and $Fe^{2+}$.

A second way of expressing the same information is to give **electrode potentials** (Table 6-8). Electrode potentials are also important in that their direct measurement sometimes provides an experimental approach to the study of oxidation–reduction reactions within cells. To measure an electrode potential it must be possible to reduce the oxidant of the couple by flow of electrons (Eq. 6-62) from an electrode surface, often of specially prepared platinum.

$$A + 2\ H^+ + 2\ e^- \rightarrow AH_2 \qquad (6\text{-}62)$$

Equation 6-62 represents a reversible reaction taking place at a single electrode. A complete electrochemical cell has two electrodes and the reaction occurring is the sum of two half-reactions. The electrode potential of a given half-reaction is obtained from the measured electromotive force of a complete cell in which one half-reaction is that of a standard **reference electrode** of known potential. Figure 6-3 indicates schematically an experimental setup for measurement of an electrode potential. The standard hydrogen electrode consists of platinum over which is bubbled hydrogen gas at one atmosphere pressure. The electrode is immersed in a solution containing hydrogen ions at unit activity ($a_{H^+} = 1$). The potential of such an electrode is conventionally taken as zero. In practice it is more likely that the reference electrode will be a calomel electrode or some other electrode that has been established experimentally as reliable and whose potential is accurately known.

The standard electrode is connected to the experimental electrode compartment by an electrolyte-filled bridge. In the experimental compartment the reaction represented by Eq. 6-62 occurs at the surface of another electrode (often platinum). The voltage difference between the two electrodes is measured with a potentiometer. The difference between the observed voltage and that of the reference electrode gives the electrode potential of the couple under investigation. It is important that the electrode reaction under study be strictly reversible. When the electromotive force (emf) of the experimental cell is balanced with the potentiometer against an external voltage source, no current flows through the cell. However, for a reversible reaction a slight change in the applied voltage will lead to current flow. The flow will be in either of the two directions, depending upon whether the applied voltage is raised or lowered.

Not all redox couples are reversible. This is especially true of organic compounds; for example, it is not possible to determine readily the electrode potential for an aldehyde–alcohol couple. In some cases, e.g., with enzymes, a readily reducible dye with a potential similar to that of the couple being measured can be added. A list of suitable dyes has been described by Dutton.[51] If the dye is able to rapidly exchange electrons with the couple being studied, it is still possible to measure the electrode potential directly. In many cases electrode potentials appearing in tables have been calculated from Gibbs energy data. The student should be able to calculate many of the potentials in Table 6-8 from Gibbs energy data from Table 6-4. If A, $H^+$, and $AH_2$ are all present at unit activity in the experimental cell, the observed potential for the half-reaction will be the **standard electrode potential** $E°$. If the emf of the hypothetical cell with the standard hydrogen electrode is positive when electron flow is in the direction indicated by the arrow in Fig. 6-3, the potential of the couple $A / AH_2$ is also taken as positive (and is often called a **reduction potential**). This is the convention used in establishing Table 6-8, but potentials of exactly the opposite sign (oxidation potentials) are used by some chemists. To avoid confusion in reading it is best to be familiar with values of one or two potentials such as those of the $O_2$–$H_2O$ and $NAD^+$–NADH couples.

When electrons flow in the external circuit the maximum amount of work that can be done per mole of electrochemical reaction ($-\Delta G$) is given by Eq. 6-63

$$\begin{aligned}\Delta G &= nEF = nE \times 96.487\ \text{kJ mol}^{-1}\ \text{V}^{-1} \\ &= -nE \times 23.061\ \text{kcal mol}^{-1}\ \text{V}^{-1}\end{aligned} \qquad (6\text{-}63)$$

where $F$ equals the number of coulombs per mole of electrons (Avogadro's number multiplied by the charge on the electron = 96,487 coulombs). $E$ is measured in volts and represents the difference of the

electrode potentials of the two half-cells. In the case of a cell using the standard hydrogen electrode, $E$ is the electrode potential of the experimental couple. The number of moles of electrons transferred in the reaction equation ($n$) is usually 1 or 2 in biochemical reactions (2 for Eq. 6-62).

Since the reactants and products need not be at unit activity, we must define the observed electrode potential $E$ as a function of $E°$ and the activities (concentrations) of A, $AH_2$, and $H^+$ (Eq. 6-64).

$$E = E° + \frac{RT}{nF} \ln \frac{[A][H^+]^2}{[AH_2]}$$

If $n = 2$

$$E = E° + 0.0296 \log \frac{[A][H^+]^2}{[AH_2]} \quad \text{volts at } 25°C \tag{6-64}$$

In the biochemical literature values of the apparent standard electrode potential at pH 7 ($E°'$) are usually tabulated instead of values of $E°$ (Table 6-8, second column). Note that $E°'$ (pH 7) for the hydrogen electrode is not zero, as it is at pH 0, but −0.414 V. Values of $E$ are related to $E°'$ by Eq. 6-64 with $[H^+]^2$

**TABLE 6-8**
**Reduction Potentials of Some Biologically Important Systems[a,b]**

| Half-reaction | $E°$ (V) | $E°'$ (pH 7) (V) | $-\Delta G'$ (pH 7) (kJ mol$^{-1}$) for oxidation by $O_2$ (per 2 electrons) |
|---|---|---|---|
| $O_2 + 4\,H^+ + 4e^- \rightarrow 2\,H_2O$ | +1.229 | +0.815 | 0.0 |
| $Fe^{3+} + e^- \rightarrow Fe^{2+}$ | 0.771 | 0.771 | 8.5 |
| $NO_3^- + 2\,H^+ + 2e^- \rightarrow NO_2^- + H_2O$ | 0.421 | 76.0 | |
| Cytochrome $f$ ($Fe^{3+}$) + $e^-$ → cytochrome $f$ ($Fe^{2+}$) | | 0.365 | 86.8 |
| $Fe(CN)_6^{3-}$ (ferricyanide) + $e^-$ → $Fe(CN)_6^{4-}$ | | 0.36 | 87.8 |
| $O_2 + 2H^+ + 2\,e^- \rightarrow H_2O_2$ | 0.709 | 0.295 | 100.3 |
| Cytochrome $a$ ($Fe^{3+}$) + $e^-$ → Cytochrome $a$ ($Fe^{2+}$) | | 0.29 | 101.3 |
| $p$-Quinone + 2 $H^+$ + 2 $e^-$ → hydroquinone | 0.699 | 0.285 | 102.3 |
| Cytochrome $c$ ($Fe^{3+}$) + $e^-$ → cytochrome $c$ ($Fe^{2+}$) | | 0.254 | 108.3 |
| Adrenodoxin ($Fe^{3+}$) + $e^-$ → adrenodoxin ($Fe^{2+}$) | | 0.15 | 128.3 |
| Cytochrome $b_2$ ($Fe^{3+}$) + $e^-$ → cytochrome $b_2$ ($Fe^{2+}$) | | 0.12 | 134.1 |
| Ubiquinone + 2 $H^+$ + 2 $e^-$ → ubiquinone $H_2$ | | 0.10 | 138.0 |
| Cytochrome $b$ ($Fe^{3+}$) + $e^-$ → cytochrome $b$ ($Fe^{2+}$) | | 0.075 | 142.8 |
| Dehydroascorbic acid + 2 $H^+$ + 2 $e^-$ → ascorbic acid | | 0.058 | 146.1 |
| $Fumarate^{2-}$ + 2 $H^+$ + 2 $e^-$ → $succinate^{2-}$ | | 0.031 | 151.3 |
| Methylene blue + 2 $H^+$ + 2 $e^-$ → leucomethylene blue (colorless) | | 0.011 | 155.2 |
| Crotonyl-CoA + 2 $H^+$ + 2 $e^-$ → butyryl-CoA | | −0.015 | 160.2 |
| Glutathione + 2 $H^+$ + 2 $e^-$ → 2-reduced glutathione | | ~ −0.10 | 176.6 |
| $Oxaloacetate^{2-}$ + 2 $H^+$ + 2 $e^-$ → $malate^{2-}$ | | −0.166 | 189.3 |
| $Pyruvate^-$ + 2 $H^+$ + 2 $e^-$ → $lactate^{1-}$ | | −0.185 | 193.0 |
| Acetaldehyde + 2 $H^+$ + 2 $e^-$ → ethanol | | −0.197 | 195.3 |
| Riboflavin + 2 $H^+$ + 2 $e^-$ → dihydroriboflavin | | −0.208 | 197.4 |
| Acetoacetyl-CoA + 2 $H^+$ + 2 $e^-$ → β-hydroxybutyryl-CoA | | −0.238 (38°C) | 203.2 |
| S + 2 $H^+$ + 2 $e^-$ → $H_2S$ | 0.14 | −0.274 | 210.2 |
| Lipoic acid + 2 $H^+$ + 2 $e^-$ → dihydrolipoic acid | | −0.29 | 213.2 |
| $NAD^+$ + $H^+$ + 2 $e^-$ → NADH | −0.113 | −0.32 | 219.0 |
| $NADP^+$ + $H^+$ + 2 $e^-$ → NADPH | | −0.324 | 219.8 |
| Ferredoxin ($Fe^{3+}$) + $e^-$ → ferredoxin ($Fe^{2+}$) (*Clostridia*) | | −0.413 | 237.0 |
| 2 $H^+$ + 2 $e^-$ → $H_2$ | 0 | −0.414 | 237.2 |
| $CO_2$ + $H^+$ + 2 $e^-$ → $formate^-$ | | −0.42 (30°C) | 238.3 |
| Ferredoxin ($Fe^{3+}$) + $e^-$ → ferredoxin ($Fe^{2+}$) (spinach) | | −0.432 | 240.6 |

[a] A compound with a more positive potential will oxidize the reduced form of a substance of lower potential with a standard free energy change $\Delta G° = -n\mathrm{F}\,\Delta E° = -n\,\Delta E° \times 96.49$ kJ mol$^{-1}$ where $n$ is the number of electrons transferred from reductant to oxidant. The temperature is 25°C unless otherwise indicated. $E°$ refers to a standard state in which the hydrogen ion activity = 1; $E°'$ refers to a standard state of pH 7, but in which all other activites are unity.

[b] The major source is Loach, P. A. (1976) in *Handbook of Biochemistry and Molecular Biology* 3rd ed. Vol. I (Fasman, G. D. ed.), pp. 122–130, CRC Press, Cleveland, Ohio.

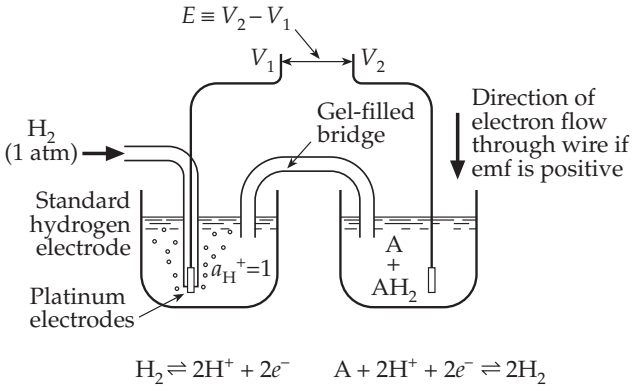
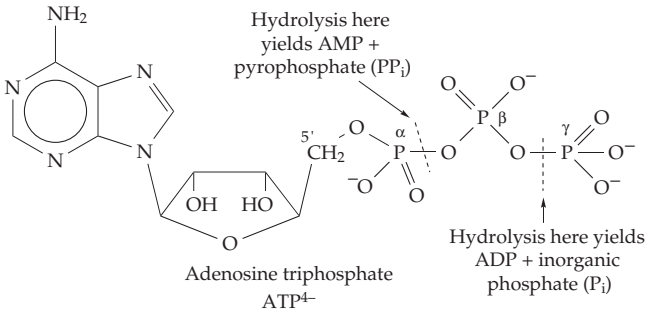



**Figure 6-3** Device for measurement of electrode potentials. The electrode reactions are indicated below each half-cell. The maximum electrical work that can be done by such a cell on its surroundings is $-\Delta G = nEF$, where $E = V_2 - V_1$ as measured by a potentiometer. If A is reduced to $AH_2$ by $H_2$, electrons will flow through an external circuit as indicated. A will be reduced in the right-hand cell. $H_2$ will be oxidized to $H^+$ in the left-hand cell. Protons will flow through the gel bridge from left to right as one of the current carriers in the internal circuit.

deleted from the numerator (since the term in log $[H^+]$ is contained in $E\,°'$). On the scale of $E\,°'$ (pH 7) the potential of the oxygen–water couple is 0.815 V, while that of the $NAD^+$–NADH couple is −0.32 V.

## D. The Adenylate System




Of central importance to the energy metabolism of all cells is the **adenylate system** which consists of adenosine 5'-triphosphate (**ATP**), adenosine 5'-diphosphate (**ADP**), and adenosine 5'-monophosphate (**AMP**) together with **inorganic phosphate** ($P_i$), **pyrophosphate** ($PP_i$), and **magnesium** ions. Remember that $P_i$ refers to the mixture of ionic forms of phosphoric acid present under experimental conditions. Between pH 4 and pH 10 this will be mainly $H_2PO_4^-$ ($pK_a = 6.8$) and $HPO_2^{2-}$. Likewise, the symbols AMP, ADP, and ATP refer to mixtures of ionic forms and $PP_i$ refers to a mixture of the ions of pyrophosphoric acid. Above pH 4.4 only $H_2P_2O_7^{2-}$ ($pK_a = 6.1$), $HP_2O_7^{3-}$ ($pK_a = 9.0$), and $P_2O_7^{4-}$ contribute appreciably to $PP_i$.

### 1. Storage and Utilization of Energy

ATP is a thermodynamically unstable molecule with respect to hydrolysis to either ADP or AMP as is indicated in the foregoing diagram. The standard Gibbs energy of hydrolysis, $\Delta G'$, of $ATP^{4-}$ to $ADP^{3-} + HPO_4^{2-}$ at pH 7 is −34.5 kJ $mol^{-1}$ and that of hydrolysis of $ADP^{3-}$ to $AMP^{2-} + HPO_4^{2-}$ is −36.3 kJ $mol^{-1}$ at 25°C (Table 6-6). The exact value of these changes in Gibbs energy depends on pH and on the concentration of $Mg^{2+}$ as is detailed in Sections B, 4, 5. Rates of reaction of components may also depend upon metal ions. Magnesium ion is especially important and complexes such as $MgATP^{2-}$ are regarded as the true substrates for many ATP-utilizing enzymes.

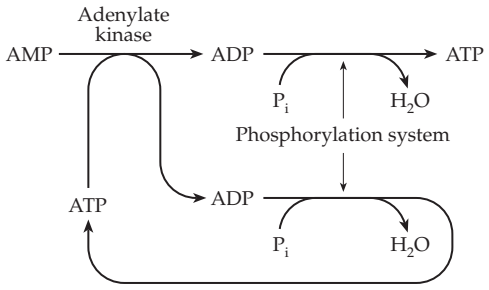
The large *Gibbs energy decreases upon hydrolysis (high group transfer potentials)* enable cells to use ATP and ADP as stores of readily available energy. Energy from the adenylate system is used for many purposes including biosynthesis, transport of ions and molecules across membranes, and for doing mechanical work. The mechanisms by which this energy is utilized are considered later (see Chapters 10, 12, and 17–19). The first step most often involves transfer of either the terminal γ-phospho group of ATP to a site on a different molecule or transfer of the entire AMP portion of the molecule onto another group. Thus the products of cleavage of ATP in these energy-utilizing processes may be either ADP + $P_i$ or AMP + $PP_i$. In the latter case the pyrophosphate is usually hydrolyzed rapidly to two molecules of $P_i$ by pyrophosphatases. This process, too, serves an essential function in the adenylate system, because it removes a product of the initial ATP cleavage reaction, shifting the overall equilibrium of the reaction sequence in the direction of the products.

The adenylate system provides the major store of rapidly available energy but the whole family of **nucleoside triphosphates** that are related to ATP in structure have similar functions. These include guanosine triphosphate (GTP), uridine triphosphate (UTP), cytidine triphosphate (CTP), and deoxyribonucleotide triphosphates (Table 5-1). These compounds are formed by successive transfer to the nucleoside monophosphate (GMP, UMP, CMP, dGMP, etc.) from two different molecules of ATP of two phospho groups. The resulting compounds are used to provide energy for a variety of specific biosynthetic processes, including synthesis of RNA and DNA. Inorganic **polyphosphates**, linear polymers of orthophosphate ($P_i$), are present in nearly all living forms.[52–54] Like ATP, they can also store

energy and in some organisms substitute for ATP in certain enzymatic reactions.

## 2. Synthesis of ATP

The **phosphate anhydride** (pyrophosphate) linkages of ATP are generated by the joining of ADP and inorganic phosphate by means of special **phosphorylation** reactions. The most important of the latter occur in the photosynthetic membranes of chloroplasts (**photosynthetic phosphorylation**) and in oxygen-utilizing membranes of bacteria and of mitochondria (**oxidative phosphorylation**). Conversion of AMP to ADP is accomplished by transfer of the terminal phospho group from an ATP molecule to AMP in a reaction catalyzed by the extremely active enzyme **adenylate kinase** (Chapter 12) which is found in all cells. The following equations indicate how one of the special phosphorylation reactions must be used twice for the conversion of one molecule of AMP into one molecule of ATP.




(6-65)

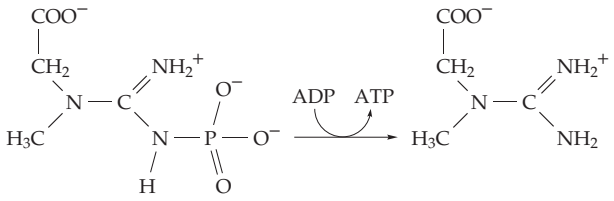
Various measures of the phosphorylating potential of the adenylate system within cells have been proposed. One measure is the product [ATP] / [ADP][$P_i$], which will be called the **phosphorylation state ratio** or $R_p$ in this book. It is also sometimes called the "phosphorylation potential." This ratio directly affects the Gibbs energy of hydrolysis of ATP, as is shown by Eq. 6-29. The value of $R_p$ may be as high as $10^4$ to $10^5$ $M^{-1}$ within cells[55] adding – 22.8 kJ $mol^{-1}$ to $\Delta G$ of hydrolysis of ATP. Another quantity that is sometimes cited is the "energy charge," the mole fraction of adenylic acid "charged" by conversion to ATP. ADP is regarded as "half-charged."[56,57]

$$\text{Energy charge} = \frac{[\text{ATP}] + \frac{1}{2}[\text{ADP}]}{[\text{ATP}] + [\text{ADP}] + [\text{AMP}]} \qquad (6\text{-}66)$$

The energy charge varies from 0 if only AMP is present to 1.0 if all of the AMP is converted to ATP. Measurements on a variety of cells and tissues show that the energy charge is usually between 0.75 and 0.90. Although it is easy to calculate its numerical value, the energy charge cannot be used in chemical equations and the idea that the energy charge of a cell plays a key role in regulation of metabolism has been challenged.[58]

## 3. Creatine Phosphate, an Energy Buffer

Although ATP provides the immediate source of energy for operating muscle, its concentration is only about 5 mM. However, muscle contains, in addition, a **phosphagen**, an *N*-phospho derivative of a guanidinium compound. In mammalian muscle, the phosphagen is **creatine phosphate**. Related compounds including **arginine phosphate** serve in various invertebrates. The group transfer potential ($-\Delta G°'$ of hydrolysis) for creatine phosphate is 43.1 kJ $mol^{-1}$. Thus, the transfer of a phospho group from creatine phosphate to ADP to form ATP (Eq. 6-67) is spontaneous with $\Delta G' = -8.6$ kJ $mol^{-1}$. Recent values under a variety of conditions have been reported by Taugue and Dobson.[59] Creatine phosphate, which is present at a concentration of 20 mM, provides a reserve of high-energy phospho groups and keeps the adenylate system of muscle buffered at a high value of $R_p$.




$\Delta G' = -8.6$ kJ $mol^{-1}$ (6-67)

## 4. Phosphorus-31 NMR

A spectacular development is the ability to observe components of the adenylate system as well as phosphocreatine and other phosphate esters in living cells by $^{31}P$ NMR.[60–61a] Spectra can be recorded on suspensions of cells[60] or of mitochondria,[62] on individual excised muscles (Fig. 6-4),[63,64] or on perfused organs[65,66] including beating rat hearts and surgically exposed animal organs.[61,61a] Metabolism can be observed in human erythrocytes[67] and even in human limbs, liver, hearts, and brain.[68–70] The method can play a valuable role in understanding human diseases.[63,67,70] As is seen in Fig. 6-4, each phosphorus atom of ATP gives a distinct resonance. The area of the $P_\gamma$ peak provides a direct measure of the ATP concentration, and

phosphocreatine (creatine-*P*) and $P_i$ can be estimated in a similar way. Knowing the concentrations of ATP, $P_i$, and creatine-*P*, the amount of ADP present, usually too low to be estimated by NMR, may be calculated.[61] Barnacle muscle contains, instead of phosphocreatine, a high concentration of phosphoarginine which has also been measured by NMR.[71]

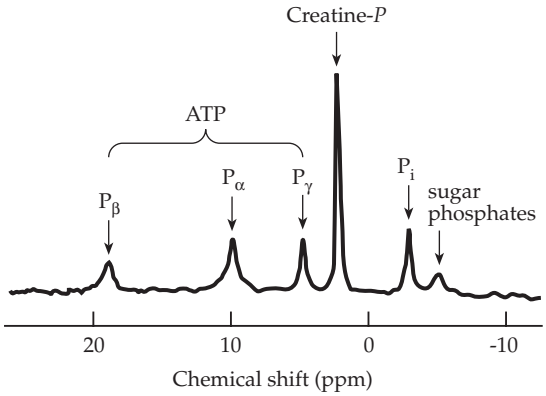



**Figure 6-4** Phosphorus-31 NMR spectrum of an excised rat muscle (*vastus lateralis*) in Ringer solution at 15°C. The spectrum represents the accumulation of 400 scans. From P. J. Seeley *et al.*[64]

Usually no ADP can be seen in NMR spectra, although there may be about 0.5 mM ADP according to chemical analysis of rapidly frozen tissues. It has been concluded that most ADP is bound to proteins. In muscle the proteins myosin and actin (Chapter 20) hold most of the ADP leaving only about 0.02 mM free.[62,72] The same conclusion has been reached on the basis of other evidence.[55]

If a perfused heart in an NMR spectrometer is deprived of oxygen, the level of phosphocreatine drops rapidly and that of $P_i$ increases while that of ATP remains relatively constant until the phosphocreatine is gone. Then it too falls and becomes undetectable after 17 min.[72] Similar changes occur when a tourniquet is placed on a human arm in the NMR spectrometer. Study of the rate and extent of recovery of the adenylate system when oxygen is readmitted is helping to provide a better means of protecting kidneys and other organs during transplantation operations. Diagnosis of circulatory ailments in human limbs by $^{31}P$ NMR may soon be routine. Use of radiofrequency coils placed on the body surface allows monitoring of the adenylate system in the heart, brain, and other tissues deep within the body.[69,73] Changes of concentration in the adenylate system and of phosphocreatine can be monitored very rapidly and evenly throughout the cardiac heartbeat cycle.[74]

## E. Complex Biochemical Equilibria

The binding of small molecules to larger ones is basic to most biological phenomena. Substrates bind to enzymes and hormones bind to receptors. Metal ions bind to ATP, to other small molecules, and to metalloproteins. Hydrogen ions bind to amino acids, peptides, nucleotides, and most macromolecules. In this section we will consider ways of describing mathematically the equilibria involved.

The strength of bonding between two particles can be expressed as a **formation constant** (or **association constant)** $K_f$. Consider the binding of a molecule X to a second molecule P, which may be a protein or some other macromolecule. If there is on the surface of P only one single binding site for X, the process can be described by Eq. 6-68 and the equilibrim constant $K_f$ by Eq. 6-69.

$$X + P \rightarrow PX \tag{6-68}$$

$$K_f = [PX] / [P][X] \tag{6-69}$$

The units for $K_f$ are liters per mole ($M^{-1}$). The constant $K_f$ is a direct measure of the strength of the binding: The higher the constant, the stronger the interaction. This fact can be expressed in an alternative way by giving the standard Gibbs energy change ($\Delta G°$) for the reaction (Eq. 6-70). The more negative $\Delta G_f°$, the stronger the binding.

$$\begin{aligned}\Delta G_f° &= -RT \ln K_f = -2.303RT \log K_f \\ &= -5.708 \log K_f \text{ kJ mol}^{-1} \text{ at } 25°C\end{aligned} \tag{6-70}$$

To avoid confusion, it is important to realize that the frequently used **dissociation constants** ($K_d$) are reciprocal association constants (Eq. 6-71). The use of

$$K_d = 1 / K_f \tag{6-71}$$

association constants and of dissociation constants is firmly entrenched in different parts of the chemical literature; be sure to keep them straight. Dissociation constants are customarily used to describe acid–base chemistry, while formation constants are more often employed to describe complexes with metal ions or associations of macromolecules (Section C). However, both types of equilibria can be described using either formation constants or dissociation constants.

Logarithms of formation constants, which are proportional to the Gibbs energies of association, are often tabulated. The logarithms of formation constants and $pK_a$ values of dissociation constants are identical (Eq. 6-72) and are a measure of the standard

$$\log K_f = -\log K_d = pK_d \tag{6-72}$$

Gibbs energy decrease in the association reaction. The difference in $\Delta G^\circ$ corresponding to a change of one unit in log $K_f$ or p$K_d$ is –5.7 kJ mol$^{-1}$, a handy number to remember.

## 1. Effects of pH on Equilibria

Compounds that contain several acidic or basic groups can exist in a number of different ionic forms, $H_nA$, $H_{n-1}A$, $H_2A$, HA, A, etc., as is indicated in the following equation, which is identical in form to Eq. 3-3.

$$H_nA \xrightarrow{K_1} H_{n-1}A \longrightarrow \cdots\cdots H_2A \xrightarrow{K_{n+1}} HA \xrightarrow{K_n} A \qquad (6\text{-}73)$$

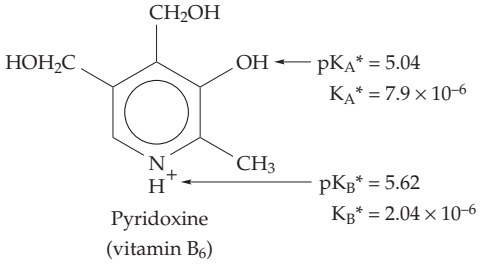
The dissociation constants $K_1$ ---- $K_n$ for a multiprotic acid $H_nA$ are defined as stepwise or **macroscopic constants** (also called molecular constants). For some compounds, e.g. alanine, the p$K_a$ values are far apart (p$K_1$ and p$K_2$ are 2.4 and 9.8, respectively). The macroscopic constants can be assigned specifically, $K_1$ to the carboxyl group and $K_2$ to the protonated amino group. At the isoelectric pH of 6.1 the alanine exists almost entirely as the dipolar ion. However, for compounds in which the macroscopic p$K_a$ values are closer together, they cannot be assigned to specific groups. We will consider some specific examples in the next section.

## 2. Microscopic Dissociation Constants and Tautomerization

A microscopic constant applies to a single site. Consider the dissociation of a simple carboxylic acid:

$$R—COOH \longrightarrow R—COO^- + H^+ \qquad (6\text{-}74)$$

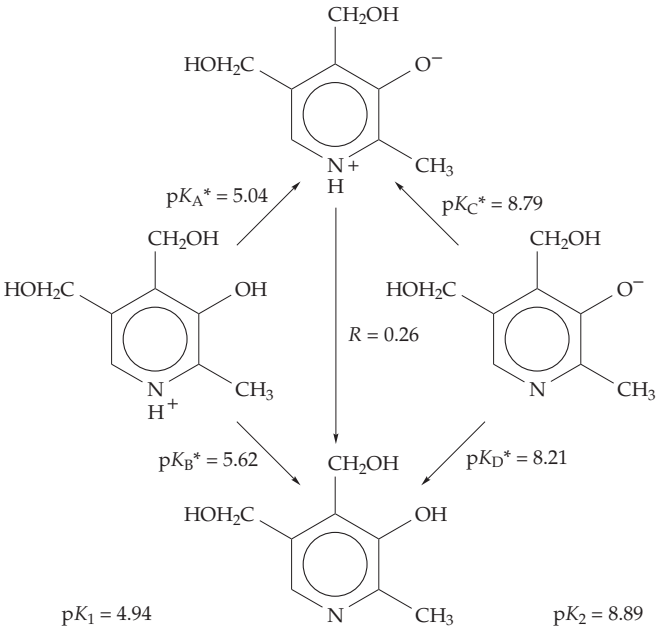
The dissociation constant is about 1.7 x 10$^{-5}$ and p$K_a$ = 4.8. Since there is only one proton, the observed dissociation constant is also the microscopic dissociation constant. Now consider the cation of pyridoxine which has two dissociable protons bound to distinctly different sites, the phenolic oxygen and the ring nitrogen.




Pyridoxine
(vitamin $B_6$)

Either of the two protons might dissociate first as the pH is raised (Eq. 6-75). However, the two microscopic dissociation constants p$K_A$* and p$K_B$* are distinctly different. The result is that at 25°C in the neutral (monoprotonated) form 80% of the molecules carry a proton on the N, while the other 20% are protonated on the less basic – O$^-$. Notice that the subscripts a and b used in this discussion do *not* refer to acidic and basic but to the individual dissociation steps shown in Eq. 6-75. Microscopic constants will always be indicated with asterisks in this discussion.

The two monoprotonated forms of pyridoxine are the tautomeric pair shown in Eq. 6-75 and whose concentrations are related by the **tautomeric ratio**, $R$ = [neutral form]/[dipolar ion], a *pH-independent equilibrium constant* with a value of 0.204/0.796 = 0.26 at 25°C.[75] Evaluation of microscopic constants for dissociation of protons from compounds containing non-identical groups depends upon measurement of the tautomeric ratio, or ratios if more than two binding sites are present. In the case of pyridoxine, a spectrophotometric method was used to estimate $R$.




(6-75)

To calculate microscopic constants from stepwise constants and tautomeric ratios, consider Eq. 6-76 in which $[HP]_A$ and $[HP]_B$ are the concentrations of the two tautomers and $K_1$ is the first **stoichiometric** or **macroscopic** dissociation constant for the diprotonated species $H_2P$.

$$K_1 = \frac{([HP]_A + [HP]_B)[H^+]}{[H_2P]} = K_A^* + K_B^*$$

$$pK_1 = 4.94;\ K_1 = 1.15 \times 10^{-5} = 9.1 \times 10^{-6} + 2.4 \times 10^{-6} \qquad (6\text{-}76)$$

We see that $K_1$ is just the sum of the two microscopic constants $K_A$* and $K_B$* for dissociation of $H_2P$ to the pair of tautomers HP(A) and HP(B). In a similar fashion it can be shown that the second stoichiometric constant $K_2$ is related to the microscopic constants $K_C$* and $K_D$* for dissociation of HP(A) and HP(B) to form P (Eq. 6-77).

$$1 / K_2 = 1 / K_C^* + 1 / K_D^* \qquad (6\text{-}77)$$

Since the tautomeric ratio $R$ equals $[HP]_B / [HP]_A$, Eqs. 6-76 and 6-77 can be rearranged to Eqs. 6-78 to 6-81. These allow the evaluation of all of the microscopic constants from the two stoichiometric constants $K_1$ and $K_2$ plus the tautomeric ratio $R$.

$$pK_A^* = pK_1 + \log (1 + R) \qquad (6\text{-}78)$$

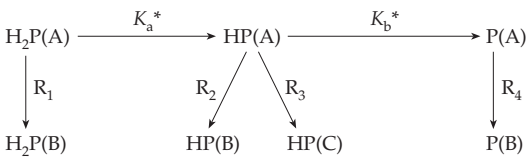
$$pK_B^* = pK_A^* - \log R \qquad (6\text{-}79)$$

$$pK_C^* = pK_2 - \log (1 + R) \qquad (6\text{-}80)$$

$$pK_D^* = pK_C^* + \log R \qquad (6\text{-}81)$$

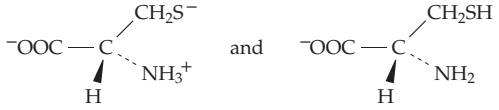
For pyridoxine $pK_1$ and $pK_2$ were determined spectrophotometrically as 4.94 and 8.89. These values, together with that of $R$ given above, were used to estimate the microscopic constants that are given in Eq. 6-74.[75] Notice that the microscopic constants of Eq. 6-74 are not all independent; if any three of the five equilibrium constants are known the other two can be calculated readily. In describing and measuring such equilibria it is desirable to select one pathway of dissociation, e.g., $H_2P \rightarrow HP(A) \rightarrow P$, and to relate the species HP(B) to it via the pH-independent constant $R$.

Often more complex situations arise in which additional tautomers or other forms arise via pH-independent reactions. These can all be related back to the reference ionic species by additional ratios $R$, which may describe equilibria for tautomerization, hydration, isomerization, etc. (Eq. 6-82).[76] In the case illustrated, only one of the ratios, namely $R_2$ or $R_3$, is likely to be a tautomerization constant because, as a rule, $H_2P$ and P will not have tautomers. Equations analogous to Eqs. 6-76 to 6-82 can be written easily to derive $K_C$*, $K_D$* and any other microscopic constants desired from the stoichiometric constants plus the ratios $R_1$ to $R_4$. While it is easy to describe tautomerism by equations such as Eqs. 6-76 and 6-82 *it is often difficult to measure the tautomeric ratios R*.[77] In favorable cases measurements of spectra of one kind or another allow their evaluation. However, because tautomerism may be extremely rapid, NMR spectra will often show only one peak for a proton present in a mixture of tautomers.




(6-82)

As was pointed out in Chapter 2, tautomerization ratios are often affected strongly by changes in solvent. Tautomerism among monoprotonated forms of cysteine, glutathione, and histidine (Eq. 2-6) has received considerable attention by biochemists[77–79] as has tautomerism in binding of protons and other small ligands to proteins.[80,81] For cysteine,[78,79] for which the following species coexist in the alkaline pH range, the distribution of the various ionic species including the two tautomers is shown in Fig. 6-5. A similar situation is met in the small protein thioredoxin (Box 15-C) which has a buried aspartate carboxylate that interacts with a nearby cysteine –SH group,[82] in papain where –SH and imidazole groups interact (Fig. 12-15),[83] and in carbohydrases where two or more carboxyl and carboxylate side chains interact (Chapter 12).[84]



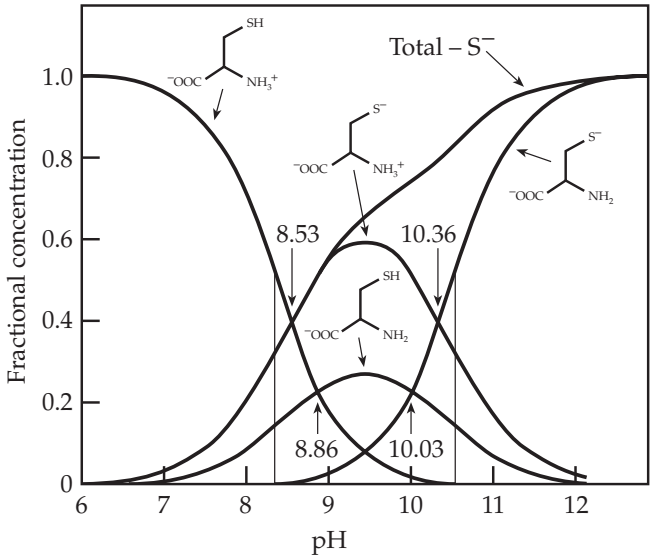
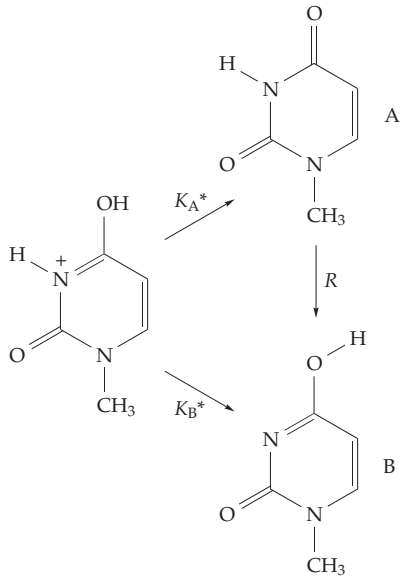



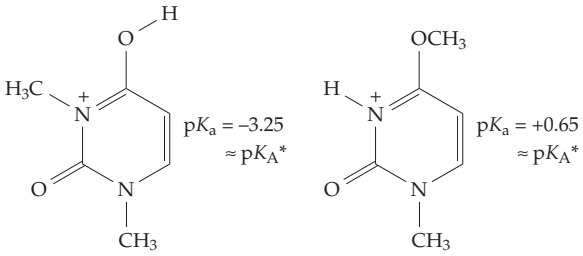

**Figure 6-5** Distribution of various ionic species of cysteine as a function of pH. The function in each ionic and tautomeric form is plotted. Microscopic $pK_a$ values are given numerically and macroscopic $pK_a$'s are indicated by the vertical lines. From Dixon and Tipton.[79]

Because it is usually difficult to measure tautomeric ratios, dissociation constants are frequently assumed identical to those of compounds in which one of the acidic groups has been modified by methylation or esterification or in some other manner to prevent dissociation of a proton. For example, the cation of 1-methyluracil can be dissociated to the two tautomers shown in Eq. 6-83.




(6-83)

The apparent p$K_a$ values were estimated for dissociation of the following two dimethylated cations which resemble that of Eq. 6-83.




It can reasonably be assumed that these p$K_a$ values approximate p$K_A$* and p$K_B$* as indicated. Thus, applying Eq. 6-85, log $R \approx -3.25 - 0.65 + 3.90$. This result indicates that tautomer A of Eq. 6-89 is overwhelmingly predominant in water.[85,86] It also suggests that, within experimental error, the p$K_a$ for 1-methyluracil will equal p$K_A$*, namely, –3.25. In fact, it is close to this (–3.40).[85]

## 3. The Binding of Metal Ions

Equilibria in the formation of complex ions with metals are treated exactly as is the binding of small molecules and ions to macromolecules.[87–89] Stepwise constants are defined for the formation of complexes containing one, two, or more ligands L bound to a central metal ion M. The binding constants $K_f$'s are usually referred to as β's as in Eq. 6-84.

$$\beta_1 = K_{f1} \quad \beta_n = K_{f1} K_{f2} \ldots K_{fn} \tag{6-84}$$

Many important questions can be asked about the binding of metal ions within living cells. For example, What fraction of a given metal ion is free and what fraction is bound to organic molecules? To what ligands is a metal bound? Since many metal ions are toxic in excess, it is clear that homeostatic mechanisms must exist. How do such mechanisms sense the free metal ion activity within cells? How does the body get rid of unwanted metal ions? Answers to all these questions depend upon the quantitative differences in the binding of metal ions to the variety of potential binding sites found within a cell.

Table 6-9 gives formation constants for 1:1 complexes of several metal ions and a number of inorganic as well as organic ligands.[89] Only the values of log $K_1$ are given when a series of stepwise constants have been established. However, in many cases two or more ligands can bind to the same metal ion. Thus for cupric ion and ammonia there are four constants.

$$Cu^{2+} + NH_3 \quad \log K_i = 4.0, 3.3, 2.7, 2.0$$

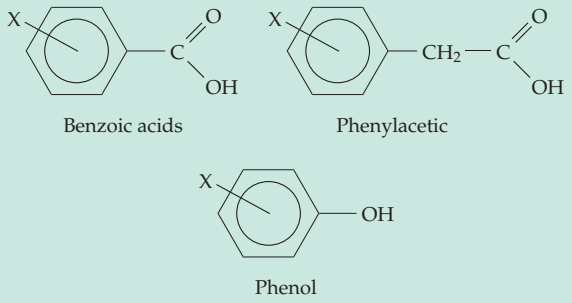
They are all separated by more than the "statistical distance," which in this case is less than the 0.6 logarithmic units for two equivalent binding sites (see Chapter 7, Section A,2). Thus, the second ligand binds less tightly than the first and **anticooperativity** in binding of successive ligands is observed. Most metal ions will also bind two or three successive amino acids. In the case of copper, whose preferred coordination number is four, two ligands may be bound. Again, a distinct anticooperativity is evident in the spread of the two constants.

$$Cu^{2+} + \text{alanine} \quad \log K_i = 8.1, 6.8$$

***Factors affecting the strength of binding of a metal in a complex.*** More basic ligands tend to bind metal ions more tightly just as they do protons. However, the strength of bonding to metal ions to a group is more nearly proportional to the **nucleophilic character** (Chapter 12) which is only partly determined by basicity to protons.

## BOX 6-C LINEAR GIBBS ENERGY RELATIONSHIPS

Organic functional groups exert characteristic electronic effects upon other groups to which they are attached. The quantitative expression of such effects can sometimes be correlated by linear Gibbs energy relationships. The best known of these is the **Hammett equation**, which deals with the transmission of electronic effects across a benzene or other aromatic ring. Consider the acid dissociation constants of three classes of compounds:




The values of p$K_a$ given in the following table have been observed for the parent compounds and for the *meta*-chloro- and the *meta*-nitro-substituted compounds.

**Values of p$K_a$ for Unsubstituted and Substituted Benzoic Acids, Phenylacetic Acids, and Phenols**

| Meta substituent | Benzoic acids* | Phenylacetic acids | Phenols |
|---|---|---|---|
| — H (parent compound) | 4.202 | 4.31 | 9.92 |
| — Cl | 3.837 | 4.13 | 9.02 |
| — $NO_2$ | 3.460 | 3.97 | 8.39 |

* For an independent set of p$K_a$ values for substituted benzoic acids see Bolton and Fleming.[a] The values of σ calculated from them are slightly different from those given here.

Since p$K_a$ is the negative logarithm of a dissociation constant, it follows from Eq. 6-30 that values of p$K_a$ are directly proportional to values of $\Delta G°$ for dissociation of protons. In the Hammett treatment differences in p$K_a$ values, rather than differences in $\Delta G°$, are considered. When a hydrogen atom in the meta position of benzoic acid is replaced by the electron-withdrawing Cl or $NO_2$, p$K_a$ is lowered, i.e., the basicity of the conjugate base of the acid is decreased. The decrease in p$K_a$ amounts to −0.365 for *m*-chlorobenzoic acid and −0.742 for *m*-nitrobenzoic acid. The Hammett treatment asserts that these changes in p$K_a$ are a measure of the electron-withdrawing power of the meta substituent.[b] Thus, the nitro group is about twice as strong as the chloro group in this respect. The numerical values of these changes in the dissociation constant of benzoic acid define the **substituent constants** σ, which are used in the Hammett equation and are given in the following table. In this equation we use the symbol p$K_0$ to represent the p$K_a$ of the unsubstituted parent compound and p$K$ to represent the p$K_a$ of the substituted molecule.

For substituted benzoic acids: $pK_0 - pK = \sigma$

The decreases in the p$K$ of phenylacetic acid occasioned by replacement of the *meta*-hydrogen with Cl or $NO_2$ are −0.18 and −0.34, respectively, substantially less than for the benzoic acids. On the other hand, for the phenols the differences amount to −0.90 and −1.53, much *greater* than for the benzoic acids. The Hammett equation asserts that for reactions such as the dissociation of protons from phenylacetic acids or from phenols, the changes in $\Delta G$ occasioned by meta substitutions are proportional to the σ values, i.e., to the changes in $\Delta G$ for the standard reaction – dissociation of a proton from benzoic acid.[b–d]

**Substituent Constants for Use in the Hammett and Taft Equations[b-d]**

| Substituent | $\sigma_m$ | $\sigma_p$ | $\sigma_p^-$ | $\sigma_p^+$ | σ* (Taft) |
|---|---|---|---|---|---|
| $-O^-$ | −0.71 | −0.52 | | | |
| $-NH_2$ | −0.16 | −0.66 | | −1.11 | 0.62 |
| −OH | 0.121 | −0.37 | | −0.85 | 1.34 |
| $-OCH_3$ | 0.115 | −0.27 | | −0.78 | 1.81 |
| $-CH_3$ | −0.069 | −0.17 | | −0.31 | 0.00 |
| $-NH-COCH_3$ | 0.21 | −0.01 | | −0.25 | |
| −H | 0 | 0 | | 0 | 0.49 |
| $-CH_2OH$ | 0.08 | 0.08 | | | 0.56 |
| $-COO^-$ | −0.10 | 0.00 | | 0.11 | −1.06 |
| $-SO_3^-$ | −0.05 | 0.09 | | 0.12 | |
| −SH | 0.25 | 0.15 | | 0.019 | 1.68 |
| $-CH_2Cl$ | | 0.18 | | | 1.05 |
| $-CONH_2$ | 0.28 | 0.36 | | | |
| −F | 0.337 | 0.06 | 0.02 | −0.07 | |
| −I | 0.352 | 0.18 | | 0.135 | |
| −Cl | 0.373 | 0.227 | | 0.114 | |
| −CHO | 0.36 | 0.22 | 0.37 | | |
| $-COCH_3$ | 0.376 | 0.502 | 0.85 | | |
| −COOH | 0.37 | 0.45 | | 0.42 | 2.08 |
| $-COOCH_3$ | 0.39 | 0.31 | | 0.49 | |
| $-SO_2NH_2$ | 0.55 | 0.62 | | 0.61 | |
| −CN | 0.56 | 0.66 | 0.89 | 0.66 | |
| $-C\equiv CH$ | | | | | 2.18 |
| $-CF_3$ | 0.43 | 0.54 | | 0.61 | 2.61 |
| $-CCl_3$ | | | | | 2.65 |
| $-NO_2$ | 0.710 | 0.778 | 1.25 | 0.790 | 4.0 |
| $-NH_3^+$ | 1.13 | 1.70 | | | 3.76 |
| N (pyridine) | 0.73 | 0.83 | | | |
| $NH^+$ (pyridine) | 2.18 | 2.42 | | 4.0 | |

**BOX 6-C (continued)**

$$\log (K / K_0) = pK_0 - pK = \rho\sigma$$

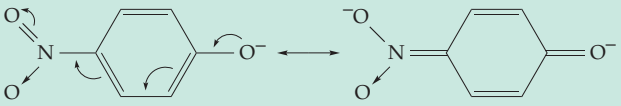
The Hammett equation

The proportionality constant ρ, which also appears in the equation, is a measure of the *sensitivity* of the reaction to the presence of electron-withdrawing or electron-donating substituents in the ring. For benzoic acid, ρ is taken as 1.00. Using the data from the accompanying table together with many other data, an average value of ρ = 0.49 has been found for phenylacetic acids. Likewise, ρ = 2.23 for phenols, and ρ = 5.7 for dissociation of protons from substituted pyridinium ions. The sensitivity to substituent changes is highest (highest ρ) in the latter case where proton dissociation is directly from an atom in the ring and is lowest when the basic center is removed farthest from the ring (phenylacetic acid).

Through knowledge of the value of ρ for a given reaction, it is possible to predict the effect of a substituent on $pK_a$ using the tabulated values of σ. In many cases the effects are additive for multisubstituted compounds.

$$pK_a = pK_0 - \rho\Sigma\sigma$$

Since substituents in *o*, *m*, and *p* positions have quantitatively different influences, different substituent constants σ are defined for each position. Moreover, since special complications arise from ortho interactions, it is customary to tabulate σ values only for meta and para positions. These are designated $\sigma_m$ and $\sigma_p$. Apparent σ values for *ortho* substituents are also available.[d] An additional complication is that certain reactions are unusually sensitive to para substituents that are able to interact by resonance directly across the ring. An example is the acid dissociation of phenols. While $\sigma_p$ for the nitro group is ordinarily 0.778, a correct prediction of the effect of the *p*-nitro group on dissociation of phenol is given only if $\sigma_p$ is taken as 1.25. This higher σ value is designated $\sigma_p^-$. The resonance in the phenolate anion giving rise to this enhanced effect of the nitro group may be indicated as follows:



(Here the curved arrows represent the direction of flow of electrons needed to convert from the one resonance structure to the other.) For similar reasons, some reactions require the use of special $\sigma_p^+$ constants for strongly electron-donating substituents such as OH which are able to interact across the ring by resonance. Thus $\sigma_p$ for the OH group is −0.37, while $\sigma_p^+$ is −0.85.[e] For the methoxyl group (—$OCH_3$) $\sigma_p = -0.27$, $\sigma_p^+ = -0.78$, and $\sigma_m = +0.12$.

The use of different kinds of substituent constants complicates the application of the Hammett equation and over 20 different sets of σ values have been proposed. A simplification is the representation of substituent constants as linear combinations of two terms, one representing "field" or "inductive" effects and the other resonance effects.[e,f]

Many other linear Gibbs energy relationships have been proposed; for example, the acid strengths of aliphatic compounds can be correlated using the "Taft polar substituent constants" σ*.

$$\log (K/K_0) = \rho^* \sigma^*$$

For example, the following give good approximations of $pK_a$ values.[d]

for R—COOH $\quad pK_a = 4.66 - 1.62\ \sigma^*$
for R—$CH_2$—COOH $\quad pK_a = 5.16 - 0.73\ \sigma^*$

While the examples chosen here concern only dissociation of protons, the Hammett equation has a much broader application. Equilibria for other types of reactions can be treated. Furthermore, since rates of reactions are related to Gibbs energies of activation, many rate constants can be correlated. For these purposes the Hammett equation can be written in the more general form in which *k* may be either an equilibrium constant or a rate constant.[b] The subscript *j* denotes the reaction under consideration and *i* the substituent influencing the reaction.

$$\log k_{ij} - \log k_{0j} = \rho_i \sigma_j$$

An example of a linear Gibbs energy relationship that is widely used in discussing mechanisms of enzymatic reactions is the Brönsted plot (Eqs. 9-90 and 9-91).

---

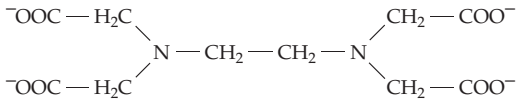
[a] Bolton, P. D., Fleming, K. A., and Hall, F. M. (1972) *J. Am. Chem. Soc.* **94**, 1033–1034
[b] Hammett, L. P. (1970) *Physical Organic Chemistry*, 2nd ed., p. 356, McGraw-Hill, New York
[c] Wells, P. R. (1963) *Chem. Rev.* **63**, 171–219
[d] Barlin, G. B., and Perrin, D. D. (1966) *Q. Rev., Chem. Soc.* **20**, 75–101
[e] Swain, C. G., and Lupton, E. C., Jr. (1968) *J. Am. Chem. Soc.* **90**, 4328–4337
[f] Hansen, L. D., and Hepler, L. G. (1972) *Can. J. Chem.* **50**, 1030–1035

**TABLE 6-9**
**Logarithms of Binding Constants for Some 1:1 Metal Complexes at 25°C[a]**

| Ligand | $H^+$ | $Mg^{2+}$ | $Ca^{2+}$ | $Mn^{2+}$ | $Cu^{2+}$ | $Zn^{2+}$ |
|---|---|---|---|---|---|---|
| Hydroxide, $OH^-$ | 14.0 | 2.5 | 1.4 | | 6.5 | 4.4 |
| Acetate$^-$ | 4.7 | ~0.65 | 0.5 | ~1.0 | 2.0 | 1.5 |
| Lactate$^-$ | 3.8 | ~1.0 | ~1.2 | 1.3 | 3.0 | 2.2 |
| Succinate$^{2-}$ | 5.2 | 1.2 | 1.2 | | 3.3 | 2 |
| $NH_3$ | 9.3 | ~0 | –0.2 | 0.8 | 4.0 | 2.4 |
| Ethylenediamine | 10.2 | 0.4 | | 2.8 | 10.8 | 6.0 |
| Glycine$^-$ | 9.6 | 2.2 | 1.4* | 2.8 | 8.2 | 5.0 |
| Glycine amide | 8.1 | | | ~1.5 | 5.3 | 3.3 |
| Alanine$^-$ | 9.7 | 2.0* | 1.2* | 3.0* | 8.1 | 4.6 |
| Aspartate$^{2-}$ | 9.6 | 2.4 | 1.6 | | 8.6 | 5.8 |
| Glycylglycine$^-$ | 8.1 | | 1.2* | 2.2* | 5.5 | 3.4 |
| Pyridine | 5.2 | | | 0.1 | 2.5 | 1 |
| Imidazole | 7.5 | | | 1.6 | 4.6 | 2.6 |
| Histidine | 9.1 | | | 3.3 | 10.2 | 6.6 |
| Adenine | 9.8 | | | | 8.9 | 6.4 |
| Citrate$^{3-}$ | 5.6 | 3.2 | 4.8 | 3.5 | ~4 | 4.7 |
| EDTA$^{4-}$ [b] | 10.2 | 8.8 | 10.6 | 13.8 | 18.7 | 16.4 |
| EGTA$^{4-}$ [c] | 9.4 | 5.3 | 10.9 | 12.2 | 17.6 | 12.6 |
| ATP$^{4-}$ | 6.5 | 4.2 | 3.8 | 4.8 | 6.1 | 4.9 |

[a] All values are for log $K_1$. Included is the highest p$K_a$ for protons. Data for amino acids are from Martell, A.E. and Smith, R.M. (1974, 1975) *Critical Stability Constants*, 3 vols., Plenum, New York. Others are from Sillén, L.G. and Martell, A.E. (1964) *Stability Constants of Metal–Ion Complexes*, Spec. Publ. No. 17, Chemical Society, London. Most constants for amino acids are for ionic strength 0.1. Some (designated by asterisks) are for zero ionic strength. The values shown for other ligands have been selected from a large number reported without examination of the original literature.

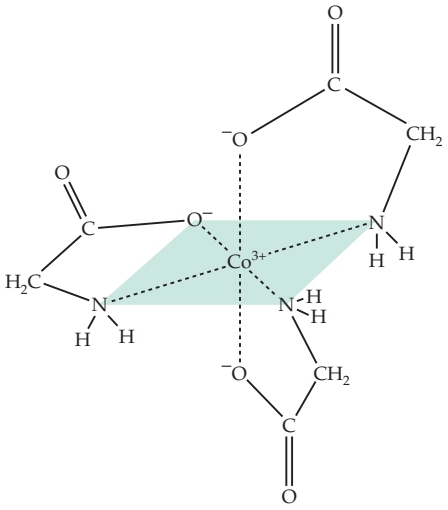
[b] Ethylenediaminetetraacetic acid (EDTA), a chelating agent widely used by biochemists for

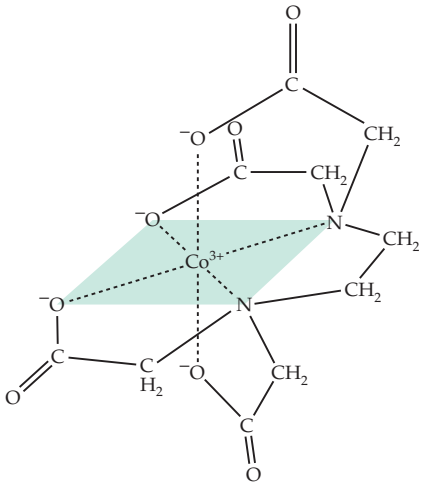


preventing unwanted reactions of metal ions. The high formation constants ensure that most metal ions remain bound to the EDTA.

[c] EGTA is similar to EDTA but has the group $-CH_2-CH_2-O-CH_2-CH_2-O-CH_2-CH_2-$ joining the two nitrogen atoms in place of $-CH_2-CH_2-$ of EDTA. Note that EGTA has a higher selectivity for $Ca^{2+}$ compared to $Mg^{2+}$ than does EDTA.

One of the most important factors in determining the affinity of organic molecules for metal ions is the **chelate effect**. The term chelate (pronounced "keel-ate") is from a Greek word meaning crab's claw. It refers to the greatly enhanced binding of metal ions resulting from the presence of two or more complexing groups in the same organic molecule. The chelate effect has been exploited by nature in the design of many important metal-binding molecules, including porphyrins (Fig. 16-5), chlorophyll (Fig. 23-20), the siderophores (Fig. 16-1), and metal-binding proteins. Structures of two **chelate complexes** are shown here. Notice from Table 6-9 that many simple compounds, such as the α-amino acids and citric acid, often form strong chelate complexes with metal ions.

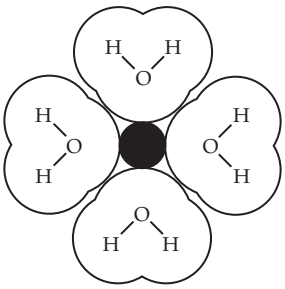




The pH of the medium always has a strong effect on metal binding. Competition with protons means that metal complexes tend to be of weak stability at low pH. Anions of carboxylic acids are completely protonated below a pH of ~4 and a metal can combine only by displacing a proton. However, at pH 7 or higher, there is no competition from protons. On the other hand, in the case of ethylenediamine, whose p$K_a$ values are 10.2 and 7.5 (Table 6-9), protons are very strong competitors at pH 7, even with a strongly complexing metal ion such as $Cu^{2+}$. At high pH there may be competition between the ligand and hydroxyl ion. At pH 7 about one-half of $Cu^{2+}$ dissolved in water is complexed as $CuOH^+$. Aluminum forms soluble complexes $Al(OH)^{2+}$, $Al(OH)_2^+$, and $Al(OH)_4^+$ (Box 12-F).

***How properties of the metal ion affect chelation.*** The **charge**, the **ionic radius** (Table 6-10), the **degree of hydration**, and the **geometry of orbitals** used in covalent bonding between metal and chelating groups all affect the formation constants of a complex. Multicharged ions form stronger complexes than do monovalent ions, which have a lower charge density.

Among ions of a given charge type (e.g., $Na^+$ vs $K^+$; $Mg^{2+}$ vs $Ca^{2+}$), the smaller ions are more strongly hydrated than are the larger ions in which the charge is dispersed over a greater surface area. Most cations, except for the large ones, have a primary hydration sphere containing about six molecules of water. Four molecules of water can be placed around the ion in one plane as shown in the following drawing of water molecules coordinated to $Mg^{2+}$.



One additional water molecule can be bound above and another below to provide six molecules in an array of **octahedral** geometry. An alternative tetrahedral arrangement of four molecules of water around the ion has been suggested for $Li^+$ and $Na^+$.[90] In either case additional solvent molecules are held in a looser secondary sphere. For instance, electrochemical transference experiments indicate a total of ~16 molecules of water around $Na^+$ and ~10 around $K^+$.

To form a chelate complex a metal ion must usually lose most of its hydration sphere. For this reason, the larger, less hydrated metal ions often bind more strongly than do the smaller, more hydrated ones. For example, $Ca^{2+}$ binds to EDTA more tightly than does $Mg^{2+}$ (Table 6-9). However, the reverse order may be observed with negatively charged ligands such as $OH^-$ in which the charge in highly concentrated. The same is true for $ATP^{4-}$, which binds $Mg^{2+}$ more tightly than $Ca^{2+}$ (Table 6-9; Section B,5).

Differences in both the charge density and the hydration of ions determine the **Hofmeister series** (lyotropic series).[91]

$$Ba^{2+} > Ca^{2+} > Mg^{2+} > Li^+ > Na^+ > K^+ > Cs^+ > NH_4^+$$

This was originally defined as the order of effectiveness in precipitating colloids or protein molecules. The ions to the left are less hydrated than those to the right. A similar series can be defined for anions. The following, a well-known sequence of the stabilities of complexes of metals of the first transition series, applies to many different types of complexes including those of the α-amino acids, as is shown in Fig. 6-11.

$$Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$$

Simple electrostatic theory based upon differences in the ionization potentials or electronegativity of the ions would predict a gradual monotonic increase in chelate stability from manganese to zinc. In fact, with nitrogen-containing ligands $Cu^{2+}$ usually forms by far the strongest complexes. Cobalt, nickel, and iron ions also show an enhanced tendency to bind to nitrogen-containing ligands. The explanation is thought to lie in the ability of the transition metals to supply *d* orbitals which can participate in covalent bond formation by accepting electrons from the ligands. Iron, copper, and cobalt are often located in the centers of nitrogen-containing structures such as the heme of our blood (iron, Fig. 16-5) and vitamin $B_{12}$ (Box 16-B).

***Metal binding sites in cells.*** Functional groups that participate in metal binding include negatively charged carboxylate $-COO^-$, thiolate $-S^-$, phenolate $-O^-$, and phosphate$^-$ as well as uncharged amino, imidazole, –OH, and the polarizable carbonyl groups of peptide and amide side chains. To which of these ligands will specific ions tend to bind? The alkali metal ions $Na^+$ and $K^+$ are mostly free within cells

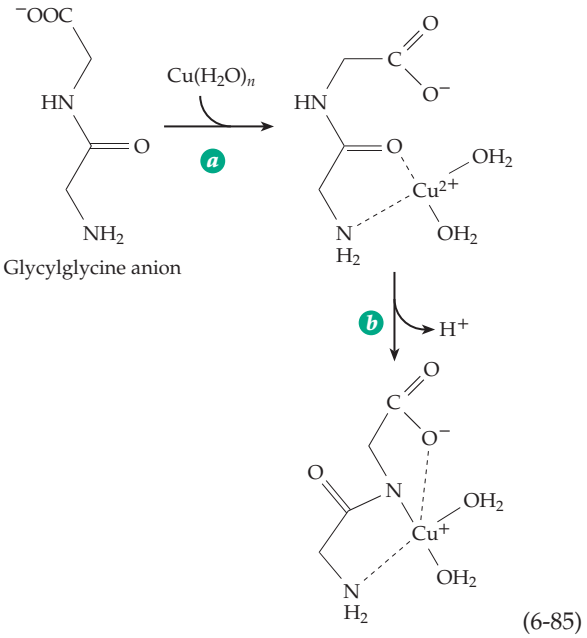
**TABLE 6-10**
**Ionic Radii in Nanometers for Some Metallic and Nonmetallic Ions[a]**

| | | | | | |
|---|---|---|---|---|---|
| | | $Mn^{2+}$ | 0.080 | | |
| $Li^+$ | 0.060 | $Fe^{2+}$ | 0.076 | $H^-$ | 0.21 |
| $Na^+$ | 0.095 | $Co^{2+}$ | 0.074 | $F^-$ | 0.136 |
| $K^+$ | 0.133 | $Ni^{2+}$ | 0.069 | $Br^-$ | 0.195 |
| $Rb^+$ | 0.148 | $Cu^{2+}$ | 0.072[b] | $I^-$ | 0.216 |
| | | $Zn^{2+}$ | 0.074 | | |
| | | $Cd^{2+}$ | 0.097 | | |
| $Be^{2+}$ | 0.031 | | | | |
| $Mg^{2+}$ | 0.065 | $Al^{3+}$ | 0.050 | | |
| $Ca^{2+}$ | 0.099 | $Fe^{3+}$ | 0.064 | | |
| $Sr^{2+}$ | 0.113 | $Mo^{4+}$ | 0.070 | | |
| $Ba^{2+}$ | 0.135 | $Mo^{6+}$ | 0.062 | | |

[a] Radii are calculated according to the method of Pauling and are taken from Cotton, F.A. and Wilkinson, G. (1972) *Advanced Inorganic Chemistry*, 3rd ed. Wiley (Interscience), New York .

[b] From Ahrens, L. H. as given by Sienko, M. J. and Plane, R. A. (1963) *Physical Inorganic Chemistry*, pp. 68–69. Benjamin, New York.

(Box 5-A) but are in part bound to defined sites in proteins. Both $Ca^{2+}$ and $Mg^{2+}$ tend to remain partially free and complexed with the numerous phosphate and carboxylate ions present in cells. However, they may find very specific tight binding sites such as that of $Mg^{2+}$ in chlorophyll (Fig. 23-20). The heavier metal ions, including those of zinc, copper, iron, and other transition metals, usually bind to nitrogen or sulfur atoms.[92] For example, small peptides react with $Cu^{2+}$ to form chelate complexes in which the peptide carbonyl oxygen binds to the metal (Eq. 6-85, step *a*).[93,94] By losing a proton the peptide NH can sometimes also function as a metal ligand (Eq. 6-85, step *b*).




(6-85)

In many **metalloproteins** the metals are found in **prosthetic groups** such as the porphyrin of heme proteins and the molybdopterin of molybdenum containing enzymes (Fig. 16-31). Very often clusters of carboxylate, imidazole, and other groups are used to create binding sites. In carboxypeptidase A (Fig. 12-16) two imidazole groups and one carboxylate of a glutamate site chain hold $Zn^{2+}$. In carbonic anhydrases, three imidazoles hold a zinc ion (Fig. 13-1), while in one kind of superoxide dismutase both $Zn^{2+}$ and $Cu^{2+}$ are bound in adjacent locations with six imidazoles and one carboxylate group participating in the binding. In contrast, $Zn^{2+}$ and $Cd^{2+}$ are bound in metallothioneins by clusters of thiol groups from cysteine side chains (Box 6-E). While $Fe^{2+}$ is bound by from four to six nitrogen atoms in heme proteins, it is also found attached to oxygen atoms of tyrosine side chains, as is shown for a transferrin in Fig. 16-2. One imidazole, one carboxylate group, and a bicarbonate ion also bind to the iron. This site is also quite satisfactory for $Al^{3+}$ (Box 12-F), which tends to occupy a fraction of the transferrin sites in blood, although it binds much more weakly than does Fe. Several other binding sites for transition metals are pictured in Chapter 16.

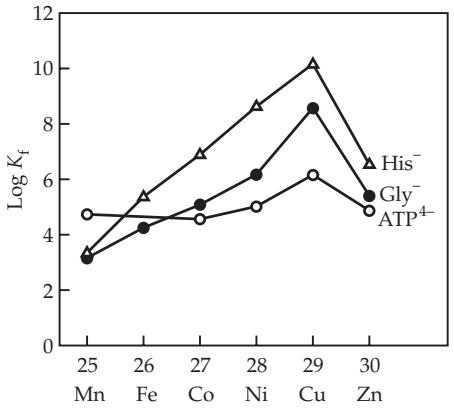



**Figure 6-6** Logarithms of formation constants of metal complexes of histidine, glycine, and ATP plotted against the atomic number of elements from manganese to zinc.

***Calcium-binding proteins.*** Much information about the functions of transition metal ions is given in Chapter 16. Here and in Box 6-D we consider calcium ions which, because of their broad distribution and range of functions, will be discussed in virtually every chapter of this book. The concentration of $Ca^{2+}$ varies greatly in different parts of a cell and also with time (Box 6-D). Many **calcium-binding proteins** participate in mediating the physiological effects of these changes and also in buffering the calcium ion concentration. These include a large family of **helix–loop–helix** or **EF-hand proteins**.[95–97] The first of these to be discovered was **parvalbumin**, a 108-residue protein from carp muscle.[96,98-99a] The structure of the pair of metal-binding sites of the protein is shown in Fig. 6-7.

Each consists of two helices that are almost perpendicular and connected by a loop that forms the $Ca^{2+}$-binding site. In the site at the left side in Fig. 6-7A the $Ca^{2+}$ is bound by four carboxylate groups from aspartate and glutamate side chains, a hydroxyl group of serine, and the residue 57 carbonyl oxygen of the peptide backbone.[99b] The same peptide group is hydrogen bonded to a carbonyl group of another segment of peptide chain near the second $Ca^{2+}$ site (to the right in Fig. 6-7A). This site contains four carboxylate ions, one of which coordinates the $Ca^{2+}$ with both oxygen atoms, and another peptide carbonyl group. By attaching itself to several different side chain groups, the metal ion can induce a substantial change in conformation from that present in the calcium-free protein.[97]

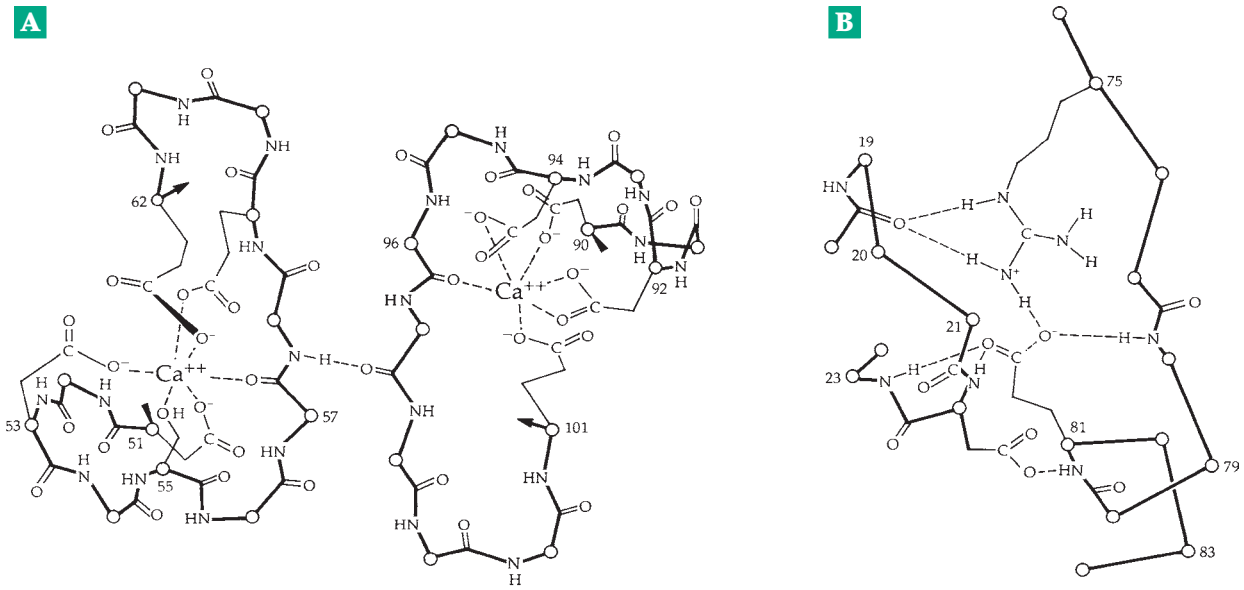



**Figure 6-7** (A) Part of the 108-residue peptide chain of the calcium-binding protein of carp muscle. The two calcium-binding loops are shown together with a hydrogen bond between them. (B) A view of the intricate network of hydrogen bonds linking two segments of the peptide chain in the interior of the molecule. Note especially the bonding of the guanidinium group from arginine-75 to the carboxylate of glutamic acid 81 and to the peptide carbonyl of residue 18. Note that the carboxylate group also interacts with two different peptide NH groups. From Kretsinger and Nockolds.[98]

Kretzinger, who discovered the structure of parvalbumin, named the $Ca^{2+}$-binding helix–loop–helix motif an **EF-hand** because it is formed by helices E and F and resembles a hand with pointer finger extended along the E helix and the thumb in the direction of the F helix, the flexed middle finger forming the $Ca^{2+}$-binding loop. Helices C and D also form a hand. The resulting pair of hands can be visualized better at the top of the calmodulin structure in Fig. 4-8. The EF-hand motif has been identified in over 1000 proteins.[100]

Parvalbumins, which are also found in other vertebrates, are high-affinity $Ca^{2+}$-buffers.[99] Additional calcium buffers with EF-hand structures are the vitamin D-induced **calbindins**. One 9-kDa calbindin is found in mammalian intestinal tissue and in skin. It has two helix–loop–helix $Ca^{2+}$-binding sites of differing affinity[101,102] that presumably function in the absorption of calcium. A 28-kDa vitamin D-dependent protein from chicken intestine contains six similar $Ca^{2+}$-binding loops.[97,103]

Another group of calcium-buffering and storage proteins with remarkable $Ca^{2+}$-binding properties are the 40- to 45-kDa **calsequestrins**, which are found in the lumen of the ER (sarcoplasmic reticulum) of skeletal muscle. Calsequestrins are not typical EF-hand proteins but have a high content of glutamate and aspartate. They bind ~50 $Ca^{2+}$ per molecule of protein with $K_d$ ~ 1 mM.[104,105] Similar proteins called **calreticulins** are found in most non-muscle cells.[106,107]

A very large number of EF-hand proteins have signaling functions. The best known of these is the 148-residue **calmodulin**, which regulates many enzymes and cellular processes (Box 6-D; also Chapter 11).[108,109] The protein, which is present in all eukaryotes, has a conserved sequence that forms two pairs of helix–loop–helix $Ca^{2+}$-binding sites that are separated by a long helix (Fig. 6-8).[108,110,111] Two of the sites bind $Ca^{2+}$ tightly and cooperatively,[110,112] with $K_d$ values in the micromolar range. Calmodulin from almost all species contains the modified amino acid ε-*N*-trimethyllysine at position 115. However, octopus calmodulin has ordinary lysine at this position and seems to function well.[108] Calmodulin's controlling functions result from $Ca^{2+}$-induced conformational changes that modify its affinity for other proteins whose activity may be increased or decreased by bound calmodulin.[109,113] A protein with a similar dumbell shape and structure is **troponin C** of skeletal muscles.[114,115] Troponin C binds to a complex of proteins that assemble on the thin actin filaments of muscle fibers and control contraction in response to changes in the calcium ion concentration (Chapter 19).[116] Other proteins that contain EF-hand motifs and are therefore responsive to $Ca^{2+}$ include **spectrin** of cell membranes,[117] **clathrin light chains** from coated vesicles,[118,119] the extracellular **osteonectin** of bones and teeth,[120] and a birch pollen antigen.[121] Another group of 17 or more small **S100 EF-hand** proteins play a variety of other roles.[122–123a] One of these, which has a high affinity for $Zn^{2+}$, has been named **psoriasin** because of its 5-fold or greater

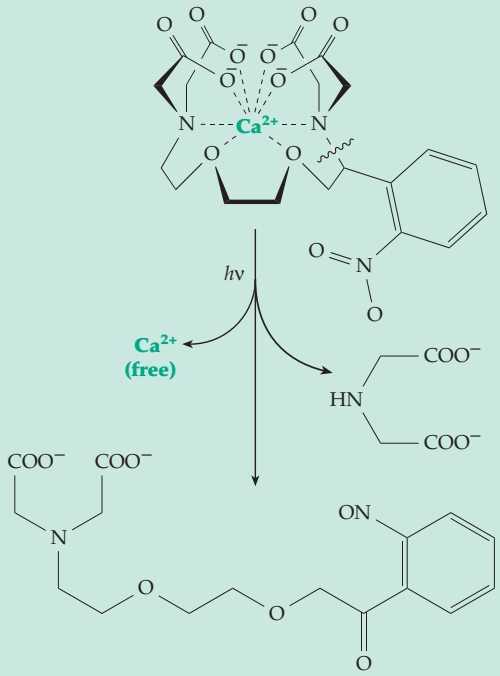
## BOX 6-D CALCIUM

The essentiality of calcium ions to living things was recognized in the last century by S. Ringer, who showed that ~1 μM $Ca^{2+}$ was needed to maintain the beat in a perfused frog heart. Later, calcium was shown essential for repair of ruptures in the cell membrane of the protozoan *Stentor* and for the motion of amebas. The animals quickly died in its absence. The role in the frog heart was traced to transmission of the nerve impulse from the nerve to the heart muscle. More recently it has been recognized that $Ca^{2+}$ is required for blood clotting and is involved in triggering many responses by cells. Calcium ions are also an integral part of many enzymes and have structural roles in proteins, carbohydrate gels, and biological membranes.

Like $Na^+$, the calcium ion is actively excluded from cells. Indeed, 99% of the calcium in the human body is present in the bones.[a–d] The blood serum concentration of $Ca^{2+}$ is ~3 mM, of which ~1.5 mM is free. The rest is chelated by proteins, carbohydrates, and other materials. Within cells the concentration of free $Ca^{2+}$ is < 1 μM and typically ~0.05–0.2 μM for unexcited cells.[d–f] However, the total intracellular $Ca^{2+}$ is considerably higher and may be in excess of 1 mM. Approximate total concentrations are: red blood cells, 20 μM; liver, 1.6 mM; and heart, 4 mM. A gradient in $[Ca^{2+}]$ of $10^3$ or more is maintained across membranes by the calcium ion pump (Chapter 8). The action of this pump is counteracted by a very slow diffusion of the external $Ca^{2+}$ back through the membrane via an $Na^+$–$Ca^{2+}$ exchange into the cells.[g]

Free $Ca^{2+}$ lacks spectroscopic properties suitable for its direct measurement at the low concentrations present in cells. However, it can be measured indirectly by the use of chelating agents that are relatively specific for $Ca^{2+}$ and which have a measurable property that changes upon calcium-binding. The fluorescent photoprotein **aequorin** (Chapter 23), which may be injected into cells or synthesized within cells from transferred genes is often employed.[h,i] Various synthetic $Ca^{2+}$ indicators also fluoresce brilliantly upon chelation.[j–n] Others contain fluorine or a suitably placed atom of $^{13}C$ which changes its NMR chemical shift upon chelation with $Ca^{2+}$.[o] These compounds may be carried into cells in the form of esters which pass through membranes but are then hydrolyzed leaving the indicators trapped in the cytoplasm.[j] Chelate compounds that bind $Ca^{2+}$ within cells and release it rapidly upon irradiation with ultraviolet light have also been developed.[p,q]

Consistent with their role in signaling, calcium ions are unevenly distributed within cells. Mitochondria, endoplasmic reticulum, Golgi, and nuclei may all take up calcium ions. Cytoplasmic $Ca^{2+}$ may sometimes be sequestered in microvesicles (**calciosomes**) or in intracellular granules[c,r,s] in which the $[Ca^{2+}]$ is ~0.5–1 mM but may reach 5–10 mM.[f] Many regulatory mechanisms exist. For example, **calcitonin** and **parathyroid hormone** interact with **vitamin D** and its metabolites in the small intestine, bones, and kidneys to control the deposition of calcium in bones, a topic considered in Box 22-C.




Calcium chelate compound that releases free calcium ion within cells upon irradiation with ultraviolet light.[p,q]

A characteristic function of $Ca^{2+}$ in living things is **activation** of various metabolic processes. This occurs when a sudden change in permeability of the plasma membrane or in the membranes of the endoplasmic reticulum (ER) allows $Ca^{2+}$ to diffuse into the cytoplasm. For example, during the contraction of muscle, the $Ca^{2+}$ concentration rises from ~0.1 to ~10 μM as a result of release from storage in the calciosomes of the ER. The calcium ions bind to **troponin C** initiating muscle contraction (Chapter 19).[t] The ER membranes of muscle are rich in a $Ca^{2+}$ pump protein[u,v] (Fig. 8-26), and in a series of calcium-binding proteins such as **calsequestrin** (see text).[f] Their combined action soon restores the $[Ca^{2+}]$ to the original low value.

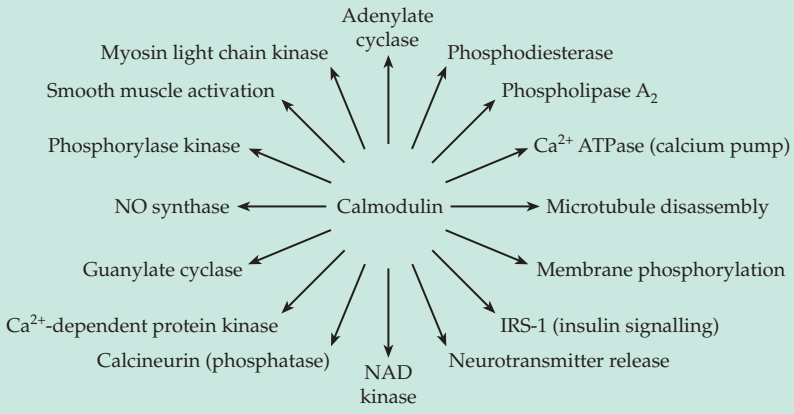
Skeletal muscle is activated by nerve impulses which induce $Ca^{2+}$ release through the action of a **voltage sensor**, a protein also known as the **dihydropyridine receptor**,[w,x] together with a **calcium release channel** known as the **ryanodine**

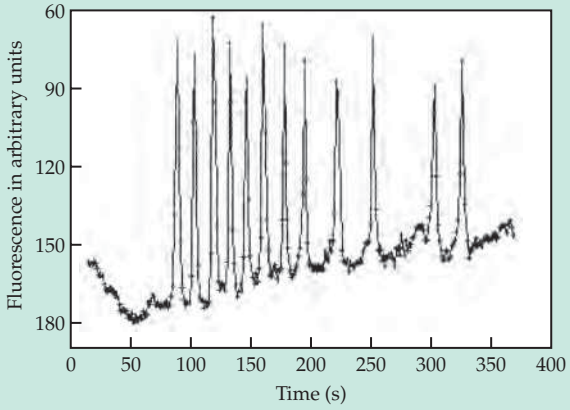
## BOX 6-D (continued)

**receptor**.[w,y–aa] Inositol trisphosphate (Fig. 11-9), cyclic ADP-ribose, as well as pH changes[bb] are involved in controlling these channels,[cc,dd] a topic discussed in Chapter 11, E. The release of $Ca^{2+}$ can be visualized using high-speed digital imaging microscopy and fluorescent [$Ca^{2+}$] indicators.[ee,ff] Release of $Ca^{2+}$ stored in sea urchin eggs is induced by cyclic ADP-ribose and by nicotinic acid adenine dinucleotide phosphate ($NAADP^+$, Eq. 15-16).[dd,gg]

In a similar manner, when a nerve impulse reaches a neuron ending (synapse) calcium ions are released into the cytoplasm and provide the trigger that causes stored neurotransmitters to be dumped into the narrow synaptic cleft that separates the endings of two communicating neurons.[y,hh] The released neurotransmitter initiates an impulse in the "postsynaptic neuron" usually again with an inflow of $Ca^{2+}$ (see Chapter 30).[e] Hormones and various other compounds often stimulate the flow of calcium ions into cells.[ii] There $Ca^{2+}$ regulates enzymes[jj,kk] (Chapter 11), microtubules, clathrin of coated vesicles (Chapter 8), $K^+$ channels in nerve membranes, and events within mitochondria,[ll,mm] in the ER,[op] and in the nucleus.[nn,oo] A substantial fraction of these responses are mediated by **calcium-binding regulatory proteins**. Among the most prominent of these is **calmodulin** (Fig. 6-8), which, upon binding of $Ca^{2+}$ activates a host of metabolic processes[pp] as indicated in the following scheme, which is modified from that of Cheung.[qq]

In many cases metabolic control by $Ca^{2+}$ is modulated by phosphorylation and dephosphorylation




Calcium oscillations observed with six cultured pancreatic β cells after a single infusion of 0.2 mM carbamoylcholine. The fluorescence intensity of the $Ca^{2+}$ indicator dye fura 2, with excitation at 380 nm, was recorded versus time. From Pretki *et al.*[ss]

or other covalent modification of proteins. Such modification may alter the affinity for $Ca^{2+}$, allowing the latter to either bind and induce a conformational change or remain unbound,[rr] without a change in [$Ca^{2+}$]. Nevertheless, the free $Ca^{2+}$ concentration within cells can change greatly and very rapidly. For example, the oscillatory change in intracellular [$Ca^{2+}$] shown above was observed in pancreatic insulin-secreting β cells responding to stimulation by the agonist carbamoylcholine. The free [$Ca^{2+}$] was evaluated from fluorescence measurements using the $Ca^{2+}$ indicator dye fura 2 (From Prentki *et al*[ss]). Oscillations in [$Ca^{2+}$] have been observed under many circumstances.[ff,oo,tt–ww] These are of particular interest because of the possible relationship to the initiation of oscillatory nerve conduction (Chapter 30).[uu] Released $Ca^{2+}$ often appears to move across cells in waves and sometimes to be released as "puffs" or "sparks."[ee,ff,oo]

[a] Bianchi, C. P. (1968) *Cell Ca*$^{++}$, Appleton, New York
[b] Means, A. R., ed. (1994) *Calcium Regulation of Cellular Function*, Lippincott-Raven, Hagerstown, Maryland
[c] Volpe, P., Krause, K.-H., Hashimoto, S., Zorzato, F., Pozzan, T., Meldolesi, J., and Lew, D. P. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1091–1095
[d] Carafoli, E. (1987) *Ann. Rev. Biochem.* **56**, 395–433
[e] Ghosh, A., and Greenberg, M. E. (1995) *Science* **268**, 239–247
[f] Meldolesi, J., and Pozzan, T. (1998) *Trends Biochem. Sci.* **23**, 10–14
[g] Khananshvili, D. (1991) *J. Biol. Chem.* **266**, 13764–13769

(continued)

**BOX 6-D (continued)**


[h] Rizzuto, R., Simpson, A. W. M., Brini, M., and Pozzan, T. (1992) *Nature (London)* **358**, 325–327
[i] Knight, M. R., Campbell, A. K., Smith, S. M., and Trewavas, A. J. (1991) *Nature (London)* **352**, 524–526
[j] Tsien, R. Y., Pozzan, T., and Rink, T. J. (1984) *Trends Biochem. Sci.* **9**, 263–266
[k] Davies, P. L., and Hew, C. L. (1990) *FASEB J.* **4**, 2460–2468
[l] Ganz, M. B., Rasmussen, J., Bollag, W. B., and Rasmussen, H. (1990) *FASEB J.* **4**, 1638–1644
[m] Lloyd, Q. P., Kuhn, M. A., and Gay, C. V. (1995) *J. Biol. Chem.* **270**, 22445–22451
[n] Etter, E., Minta, A., Poenie, M., and Fay, F. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5368–5373
[o] Robitaille, P.-M. L., and Jiang, Z. (1992) *Biochemistry* **31**, 12585–12591
[p] Ellis-Davies, G. C. R., and Kaplan, J. H. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 187–191
[q] Ellis-Davies, G. C. R., Kaplan, J. H., and Barsotti, R. J. (1996) *Biophys. J.* **70**, 1006–1016
[r] Burgoyne, R. D., and Cheek, T. R. (1991) *Trends Biochem. Sci.* **16**, 319–320
[s] Golovina, V. A., and Blaustein, M. P. (1997) *Science* **275**, 1643–1648
[t] Gagné, S. M., Li, M. X., and Sykes, B. D. (1997) *Biochemistry* **36**, 4386–4392
[u] MacLennan, D. H., Rice, W. J., and Green, N. M. (1997) *J. Biol. Chem.* **272**, 28815–28818
[v] Yonekura, K., Stokes, D. L., Sasabe, H., and Toyoshima, C. (1997) *Biophys. J.* **72**, 997–1005
[w] Yano, M., El-Hayek, R., and Ikemoto, N. (1995) *J. Biol. Chem.* **270**, 3017–3021
[x] Mitterdorfer, J., Sinnegger, M. J., Grabner, M., Striessnig, J., and Glossmann, H. (1995) *Biochemistry* **34**, 9350–9355
[y] McPherson, P. S., and Campbell, K. P. (1993) *J. Biol. Chem.* **268**, 13765–13768
[z] Sakube, Y., Ando, H., and Kagawa, H. (1997) *J. Mol. Biol.* **267**, 849–864
[aa] Zorzato, F., Fujii, J., Otsu, K., Phillips, M., Green, N. M., Lai, F. A., Meissner, G., and MacLennan, D. H. (1990) *J. Biol. Chem.* **265**, 2244–2256
[bb] Donoso, P., Beltrán, M., and Hidalgo, C. (1996) *Biochemistry* **35**, 13419–13425
[cc] Lee, M. G., Xu, X., Zeng, W., Diaz, J., Wojcikiewicz, R. J. H., Kuo, T. H., Wuytack, F., Racymaekers, L., and Muallem, S. (1997) *J. Biol. Chem.* **272**, 15765–15770
[dd] Graeff, R. M., Franco, L., De Flora, A., and Lee, H. C. (1998) *J. Biol. Chem.* **273**, 118–125
[ee] Isenberg, G., Etter, E. F., Wendt-Gallitelli, M.-F., Schiefer, A., Carrington, W. A., Tuft, R. A., and Fay, F. S. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5413–5418
[ff] Horne, J. H., and Meyer, T. (1997) *Science* **276**, 1690–1693
[gg] Lee, H. C., and Aarhus, R. (1997) *J. Biol. Chem.* **272**, 20378–20383
[hh] Miller, R. J. (1992) *J. Biol. Chem.* **267**, 1403–1406
[ii] Quitterer, U., Schröder, C., Müller-Esterl, W., and Rehm, H. (1995) *J. Biol. Chem.* **270**, 1992–1999
[jj] Suzuki, K. (1987) *Trends Biochem. Sci.* **12**, 103–105
[kk] Martin, B. L., and Graves, D. J. (1986) *J. Biol. Chem.* **261**, 14545–14550
[ll] Moudy, A. M., Handran, S. D., Goldberg, M. P., Ruffin, N., Karl, I., Kranz-Eble, P., DeVivo, D. C., and Rothman, S. M. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 729–733
[mm] Kasumi, T., Tsumuraya, Y., Brewer, C. F., Kersters-Hilderson, H., Claeyssens, M., and Hehre, E. J. (1987) *Biochemistry* **26**, 3010–3016
[nn] Santella, L., and Carafoli, E. (1997) *FASEB J.* **11**, 1091–1109
[oo] Lipp, P., Thomas, D., Berridge, M. J., and Bootman, M. D. (1997) *EMBO J.* **16**, 7166–7173
[op] Corbett, E. F., and Michalak, M. (2000) *Trends Biochem. Sci.* **25**, 307–311
[pp] Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A. (1992) *Science* **256**, 632–638
[qq] Cheung, W. Y. (1980) *Science* **207**, 19–27
[rr] Geisow, M. J. (1987) *Trends Biochem. Sci.* **12**, 120–121
[ss] Prentki, M., Glennon, M. C., Thomas, A. P., Morris, R. L., Matschinsky, F. M., and Corkey, B. E. (1988) *J. Biol. Chem.* **263**, 11044–11047
[tt] Tang, Y., and Othmer, H. G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7869–7873
[uu] Putney, J. W., Jr. (1998) *Science* **279**, 191–192
[vv] Berridge, M. J. (1997) *Nature (London)* **386**, 759–760
[ww] Hajnóczky, G., and Thomas, A. P. (1997) *EMBO J.* **16**, 3533–3543


increase above the normal level in keratinocytes of patients with the skin disease psoriasis.[123b]

Calcium ions are often involved in holding negatively charged groups together, for example, in the binding of proteins to phospholipid membranes. Among these membrane-associated proteins are the vitamin K-dependent proteins, all of which contain several residues of the chelating amino acid γ-carboxyglutamate (Chapter 15) at their calcium-binding sites.[124,125] Many of these are involved in the clotting of blood (Chapter 12). Some of the membrane-binding proteins called **annexins** (Chapter 8) are also $Ca^{2+}$-dependent ion channels.[126,127] Other $Ca^{2+}$-requiring lipid-binding proteins include the lipocortins, calpactins, and calelectrins.[128] Cadherins bind $Ca^{2+}$ and help provide cohesion between cells.[129] Many proteins contain bound $Ca^{2+}$ in precisely defined sites where it plays a structural role. These include the galactose-binding protein of bacterial transport and chemotaxis (Fig. 4-18)[130] and α-lactalbumin of milk. Although α-lactalbumin does not contain the helix–loop–helix pattern, its $Ca^{2+}$-binding site does consist of three carboxylate groups and two backbone carbonyl oxygen atoms.[131] The α-amylases, thermolysin, and staphylococcal nuclease (Chapter 12), and the lectin favin (Fig. 2-15) all contain bound $Ca^{2+}$. Calcium ions also bind to anionic groups in carbohydrates, e.g., to the sulfate groups in carageenin gels (Chapter 4) where they provide structural stability.

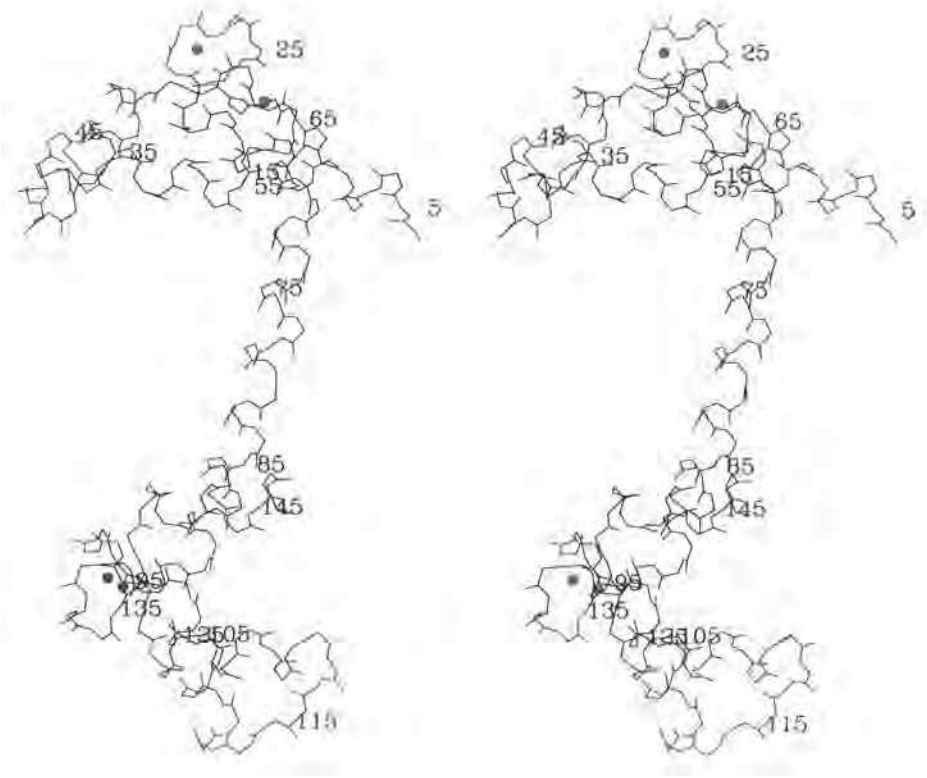
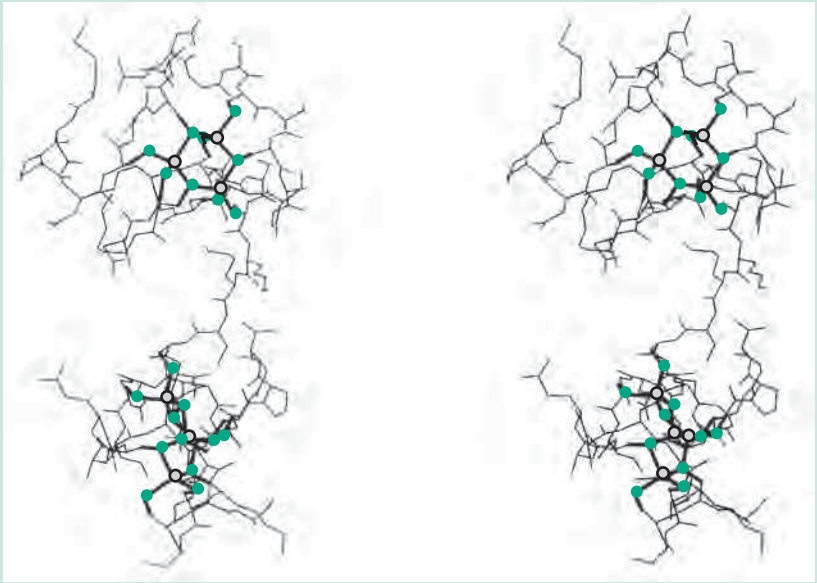



**Figure 6-8** Stereoscopic backbone trace of a 148-residue recombinant calmodulin. The two helix–turn–helix (EF-hand) loops and their bound $Ca^{2+}$ (as concentric circles) are at the top and the two near the C terminus are at the bottom. The long central helix, seen in this crystal structure, may undergo conformational changes during the functioning of this $Ca^{2+}$-sensing molecule.[132] From Chattopadhyaya *et al.*[111] Courtesy of F. A. Quiocho.

## BOX 6-E METALLOTHIONEINS

If animals ingest excessive amounts of Zn(II), Cd (II), Hg(II), or Cu(I) their livers and kidneys accumulate these metals as complexes of proteins called metallothioneins.[a–e] In mammals at least three related genes encode these metal-binding proteins. The best known, metallothionein II, has a highly conserved 61-residue sequence containing 20 cysteine residues and no aromatic residues. The protein is organized into two domains, each able to bind a cluster of metal ions via thiolate side chains. The three-dimensional structure of rat liver metallothionein containing five $Cd^{2+}$ and two $Zn^{2+}$ ions is shown in the accompanying stereoscopic diagram.[f] The 61 alpha carbons, the beta carbon and sulfur atoms (green) of cysteine residues and the bound metal ions are indicated. The N-terminal domain (residues 1–29) contains one $Cd^{2+}$ and two $Zn^2$ and nine cysteine sulfurs which bind the metal ions. Three of the sulfur atoms form bridges between pairs of metals. The second cluster contains four $Cd^{2+}$ held by 11 cysteine sulfur atoms, five of which bridge between pairs of metals. All of the metal ions are tetrahedrally coordinated. The polypeptide chains of metallothioneins consist predominantly of beta turns.[g] Important techniques in the study of these proteins include $^{113}Cd$ NMR,[h,i]



Metallothionein containing bound $Cd^{2+}$ and $Zn^{2+}$. From Robbins *et al.*[f]

## BOX 6-E METALLOTHIONEINS (continued)

spectroscopic methods,[j,k] and X-ray absorption.[l]

Transcription of metallothionein genes is induced by metal ions, and toxic metals such as Cd and Hg accumulate as metallothionein complexes, suggesting that one function is to protect against metal toxicity.[m] However, synthesis is also induced by glucocorticoid hormones,[n] and accumulation of a high concentration of copper and zinc in fetal metallothionein suggests a role in storage of these essential metals.[c] Another metal that binds to metallothioneins is Au(I),[o] which is widely used in thiolate salts as a chemotherapeutic agent for rheumatoid arthritis.

Metallothioneins are also found in insects, lower invertebrates, and even in bacteria.[i,p] Nevertheless, there are other metal-binding proteins.[q,r] For example, albacore tuna contain a 66-kDa glycoprotein that contains eight mole percent histidine and binds three $Zn^{2+}$, each by a cluster of three His.[s] Plants and fungi contain **phytochelins**, peptides consisting largely of repeated γ-glutamylcysteine units.[a,t,u] These appear to protect plants against toxicity of cadmium in the same manner as do the metallothioneins in our bodies.


[a] Mehra, R. K., Garey, J. R., Butt, T. R., Gray, W. R., and Winge, D. R. (1989) *J. Biol. Chem.* **264**, 19747–19753

[b] Kägi, J. H. R., and Kojima, Y., eds. (1987) *Metallothionein II*, Birkhäuser, Basel, Berlin

[c] Fischer, E. H., and Davie, E. W. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 3333–3334

[d] Kägi, J. H. R., and Schäffer, A. (1988) *Biochemistry* **27**, 8509–8515

[e] Hamer, D. H. (1986) *Ann. Rev. Biochem.* **55**, 913–951

[f] Robbins, A. H., McRee, D. E., Williamson, M., Collett, S. A., Xuong, N. H., Furey, W. F., Wang, B. C., and Stout, C. D. (1991) *J. Mol. Biol.* **221**, 1269–1293

[g] Pande, J., Pande, C., Gilg, D., Vasák, M., Callender, R., and Kägi, J. H. R. (1986) *Biochemistry* **25**, 5526–5532

[h] Cismowski, M. J., Narula, S. S., Armitage, I. M., Chernaik, M. L., and Huang, P. C. (1991) *J. Biol. Chem.* **266**, 24390–24397

[i] Vasák, M., Hawkes, G. E., Nicholson, J. K., and Sadler, P. J. (1985) *Biochemistry* **24**, 740–747

[j] Willner, H., Vasák, M., and Kägi, J. H. R. (1987) *Biochemistry* **26**, 6287–6292

[k] Lu, W., and Stillman, M. J. (1993) *J. Am. Chem. Soc.* **115**, 3291–3299

[l] George, G. N., Byrd, J., and Winge, D. R. (1988) *J. Biol. Chem.* **263**, 8199–8203

[m] Sadhu, C., and Gehamu, L. (1988) *J. Biol. Chem.* **263**, 2679–2684

[n] Karin, M., Haslinger, A., Holtgreve, H., Richards, R. I., Krauter, P., Westphal, H. M., and Beato, M. (1984) *Nature (London)* **308**, 513–519

[o] Laib, J. E., Shaw, C. F., III, Petering, D. H., Eidsness, M. K., Elder, R. C., and Garvey, J. S. (1985) *Biochemistry* **24**, 1977–1986

[p] Freedman, J. H., Slice, L. W., Dixon, D., Fire, A., and Rubin, C. S. (1993) *J. Biol. Chem.* **268**, 2554–2564

[q] Waalkes, M. P., and Peratoni, A. (1986) *J. Biol. Chem.* **261**, 13097–13103

[r] Willuhn, J., Schmitt-Wrede, H. P., Greven, H., and Wunderlich, F. (1994) *J. Biol. Chem.* **269**, 24688–24691

[s] Dyke, B., Hegenauer, J., Saltman, P., and Laurs, R. M. (1987) *Biochemistry* **26**, 3228–3234

[t] Reese, R. N., and Winge, D. R. (1988) *J. Biol. Chem.* **263**, 12832–12835

[u] Clemens, S., Kim, E. J., Neumann, D., and Schroeder, J. I. (1999) *EMBO J.* **18**, 3325–3333

## References


1. Lewis, G. N., and Randall, M. (1961) *Thermodynamics and the Free Energy of Chemical Substances*, 2nd ed., McGraw-Hill, New York
2. Mahan, B. H. (1963) *Elementary Chemical Thermodynamics*, Benjamin, New York
3. Everdell, M. H. (1965) *Introduction to Chemical Thermodynamics*, Norton, New York
4. Dickerson, R. E. (1969) *Molecular Thermodynamics*, Benjamin, New York
5. Jones, M. N., ed. (1988) *Biochemical Thermodynamics*, 2nd ed., Elsevier, Amsterdam
6. Hinz, H.-J., ed. (1986) *Thermodynamic Data for Biochemistry and Biotechnology*, Springer-Verlag, Berlin
7. Gutfreund, H., and Edsall, J. T. (1983) *Biothermo-dynamics*, Wiley, New York
8. Harold, F. M. (1986) *The Vital Force: A Study of Bioenergetics*, Freeman, San Francisco, California
9. Garby, L., and Larsen, P. S. (1995) *Bioenergetics; Its Thermodynamic Foundations*, Cambridge Univ. Press, London and New York
10. Jou, D., and Llebot, J. E., eds. (1990) *Introduction to the Thermodynamics of Biological Processes*, Prentice Hall, Englewood Cliffs, New Jersey

10a. Wadsö, I. (1985) *Eur. J. Biochem.* **153**, 429–434

11. Sturtevant, J. M. (1971) in *Techniques of Chemistry*, Vol. I, Part V (Weissberger, A., ed), pp. 347–425, Wiley, New York
12. Brown, H. D., ed. (1969) *Biochemical Microcalorimetry*, Academic Press, New York
13. Peters, K. S., and Snyder, G. J. (1988) *Science* **241**, 1053–1057
14. White, A., Handler, P., and Smith, E. L. (1968) *Principles of Biochemistry*, 4th ed., McGraw-Hill, New York (pp. 291–301)
15. Mosteller, R. D. (1988) *N. Engl. J. Med.* **318**, 1130
16. Donata, K., and Hegsted, D. M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4866–4870
17. Ben-Naim, A. (1992) *Statistical Thermodynamics for Chemists and Biochemists*, Plenum, New York
18. Donovan, J. W. (1984) *Trends Biochem. Sci.* **9**, 340–344
19. Tanaka, A., Flanagan, J., and Sturtevant, J. M. (1993) *Protein Sci.* **2**, 567–576
20. Pace, C. N., and Laurents, D. V. (1989) *Biochemistry* **28**, 2520–2525
21. Interunion Commission on Biothermodynamics. (1976) *J. Biol. Chem.* **251**, 6879–6885
22. Phillips, R. C., George, P., and Rutman, R. J. (1963) *Biochemistry* **2**, 501–508
23. Liu, Y., and Sturtevant, J. M. (1995) *Protein Sci.* **4**, 2559–2561
24. Halling, P. J. (1989) *Trends Biochem. Sci.* **14**, 317–318
25. Örstan, A. (1990) *Trends Biochem. Sci.* **15**, 137–138
26. Nicolis, G., and Prigogine, I. (1977) *Self-organization in Non-equilibrium Systems: From Dissipative Structure to Order Through Fluctuations*, Wiley-Interscience, New York
27. Procaccia, I., and Ross, J. (1977) *Science* **198**, 716–717
28. Morowitz, H. J. (1968) *Energy Flow in Biology*, Academic Press, New York
29. Caplan, S. R., and Essig, A. (1983) *Bioenergetics and Linear Nonequilibrium Thermodynamics/ The Steady State*, Harvard Univ. Press, Cambridge, Massachusetts
30. Peacocke, A. R. (1983) *An Introduction to the Physical Chemistry of Biological Organization*, Clarendon Press, Oxford
31. Blumenfeld, L. A. (1983) *Physics of Bioenergetic Processes*, Springer-Verlag, Berlin
32. Katchalsky, A., and Curran, P. F. (1965) *Non-equilibrium Thermodynamics in Biophysics*, Harvard Univ. Press, Cambridge, Massachusetts
33. Prigogine, I. (1967) *Introduction to Thermodynamics of Irreversible Processes*, 3rd ed., Wiley, New York
34. Coveney, P. V. (1988) *Nature (London)* **333**, 409–415
35. Wilson, D. F. (1982) *Trends Biochem. Sci.* **7**, 275–278
36. Westerhoff, H. V. (1982) *Trends Biochem. Sci.* **7**, 275–279
37. Guynn, R. W., and Veech, R. L. (1973) *J. Biol. Chem.* **248**, 6966–6972
38. Alberty, R. A. (1969) *J. Biol. Chem.* **244**, 3290–3302
39. Phillips, R. C., George, P., and Rutman, R. J. (1969) *J. Biol. Chem.* **244**, 3330–3342
40. George, P., Phillips, R. C., and Rutman, R. J. (1963) *Biochemistry* **2**, 508–512
41. Purich, D. L., and Fromm, H. J. (1972) *Curr. Top. Cell. Regul.* **6**, 131–167
42. Jencks, W. P. (1968) in *Handbook of Biochemistry* (Sober, H. A., ed), pp. J-148, CRC, Cleveland, Ohio
43. Bassham, J. A., and Krause, G. H. (1969) *Biochim. Biophys. Acta.* **189**, 207–221
44. Rosing, J., and Slater, E. C. (1972) *Biochim. Biophys. Acta.* **267**, 275–290
45. Alberty, R. A. (1992) *Biophysical Chem.* **42**, 117–131
46. Alberty, R. A., and Cornish-Bowden, A. (1993) *Trends Biochem. Sci.* **18**, 288–291
47. Alberty, R. A., and Goldberg, R. N. (1992) *Biochemistry* **31**, 10610–10615
48. Alberty, R. A. (1994) *Pure & Appl. Chem.* **66**, 1641–1666
49. Stokes, G. B. (1988) *Trends Biochem. Sci.* **13**, 422–424
50. Mavrovouniotis, M. L. (1991) *J. Biol. Chem.* **266**, 14440–14445
51. Dutton, P. L. (1971) *Biochim. Biophys. Acta.* **226**, 63–80
52. Pepin, C. A., and Wood, H. G. (1986) *J. Biol. Chem.* **261**, 4476–4480
53. Wood, H. G., and Clark, J. E. (1988) *Ann. Rev. Biochem.* **57**, 235–260
54. Crooke, E., Akiyama, M., Rao, N. N., and Kornberg, A. (1994) *J. Biol. Chem.* **269**, 6290–6295
55. Veech, R. L., Lawson, J. W. R., Cornell, N. W., and Krebs, H. A. (1979) *J. Biol. Chem.* **254**, 6538–6547
56. Swedes, J. S., Sedo, R. J., and Atkinson, D. E. (1975) *J. Biol. Chem.* **250**, 6930–6938
57. Chapman, A. G., Fall, L., and Atkinson, D. E. (1971) *J. Bacteriol.* **108**, 1072–1086
58. Purich, D. L., and Fromm, H. J. (1973) *J. Biol. Chem.* **248**, 461–466
59. Teague, W. E., Jr., and Dobson, G. P. (1992) *J. Biol. Chem.* **267**, 14084–14093
60. Cohen, S. M., Ogawa, S., Rottenberg, H., Glynn, P., Yamane, T., Brown, T. R., and Shulman, R. G. (1978) *Nature (London)* **273**, 554–556

60a. Ugurbil, K., Kingsley-Hickman, P. B., Sako, E. Y., Zimmer, S., Mohanakrishnan, P., Robitaille, P. M. L., Thoma, W. J., Johnson, A., Foker, J. E., and From, A. H. L. (1987) *Ann. N.Y. Acad. Sci.* **508**, 265–286

61. Brosnan, M. J., Chen, L., Van Dyke, T. A., and Koretsky, A. P. (1990) *J. Biol. Chem.* **265**, 20849–20855

61a. Saupe, K. W., Spindler, M., Hopkins, J. C. A., Shen, W., and Ingwall, J. S. (2000) *J. Biol. Chem.* **275**, 19742–19746

62. Hutson, S. M., Williams, G. D., Berkich, D. A., LaNoue, K. F., and Briggs, R. W. (1992) *Biochemistry* **31**, 1322–1330
63. Burt, C. T., Glonek, T., and Bárány, M. (1977) *Science* **195**, 145–149
64. Seeley, P. J., Sehr, P. A., Gadian, D. G., Garlick, P. B., and Radda, G. K. (1977) in *NMR in Biology* (Dwek, R. A., Campbell, I. D., Richards, R. E., and Williams, R. J. P., eds), pp. 247–275, Academic Press, London
65. Masson, S., and Quistorff, B. (1992) *Biochemistry* **31**, 7488–7493
66. Jeffrey, F. M., Storey, C. J., Nunnally, R. L., and Malloy, C. R. (1989) *Biochemistry* **28**, 5323–5326
67. Kagimoto, T., Higaki, T., Nagata, K., Morino, Y., and Takatsuki, K. (1989) *NMR in Biomedicine* **2**, 93–97
68. Park, J. H., Brown, R. L., Park, C. R., Cohn, M., and Chance, B. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8780–8784
69. Radda, G. K. (1992) *FASEB J.* **6**, 3032–3038
70. Radda, G. K. (1986) *Science* **233**, 640–645
71. Burt, C. T. (1985) *Trends Biochem. Sci.* **10**, 404–406
72. Grove, T. H., Ackerman, J. J. H., Radda, G. K., and Bore, P. J. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 299–302
73. Bottomley, P. A. (1985) *Science* **229**, 769–772
74. Fossel, E. T., Morgan, H. E., and Ingwall, J. S. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3654–3658
75. Metzler, D. E., Harris, C. M., Johnson, R. J., Siano, D. B., and Thomson, J. A. (1973) *Biochemistry* **12**, 5377–5392
76. Johnson, R. J., and Metzler, D. E. (1970) *Methods Enzymol.* **18**, 433–471
77. Dixon, H. B. F. (1992) *Essays in Biochemistry* **27**, 161–176
78. Vander Jagt, D. L., Hansen, L. D., Lewis, E. A., and Han, L. B. (1972) *Arch. Biochem. Biophys.* **153**, 55–61
79. Dixon, H. B. F., and Tipton, K. F. (1973) *Biochem. J.* **133**, 837–842
80. Klotz, I. M., and Hunston, D. L. (1975) *J. Biol. Chem.* **250**, 3001–3009
81. Weber, G. (1975) *Adv. Prot. Chem.* **29**, 1–83
82. Chivers, P. T., Prehoda, K. E., Volkman, B. F., Kim, B.-M., Markley, J. L., and Raines, R. T. (1997) *Biochemistry* **36**, 14985–14991
83. Johnson, F. A., Lewis, S. D., and Shafer, J. A. (1981) *Biochemistry* **20**, 44–48
84. McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Körner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., and Withers, S. G. (1996) *Biochemistry* **35**, 9958–9966
85. Katritzky, A. R., and Waring, A. (1962) *J. Chem. Soc.*, 1540–1548
86. Katritzky, A. R., and Karelson, M. (1991) *J. Am. Chem. Soc.* **113**, 1561–1566
87. Bjerrum, J. (1941) *Metal Amine Formation in Aqueous Solution*, Haase & Son, Copenhagen
88. Eichhorn, G. L., ed. (1973) *Inorganic Biochemistry*, 2 vols, Elsevier, Amsterdam
89. Martell, A. E., and Smith, R. M. (1975) *Critical Stability Constants*, Plenum, New York
90. Michaelian, K. H., and Moskovits, M. (1978) *Nature (London)* **273**, 135–136
91. Lewsin, S. (1974) *Displacement of Water and its Control in Biochemical Reactions*, Academic Press, London
92. Karlin, S., Zhu, Z.-Y., and Karlin, K. D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14225–14230
93. Freeman, H. C., Healy, M. J., and Scudder, M. L. (1977) *J. Biol. Chem.* **252**, 8840–8847
94. Torrado, A., Walkup, G. K., and Imperiali, B. (1998) *J. Am. Chem. Soc.* **120**, 609–610
95. Nakayama, S., and Kretsinger, R. H. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 473–507

## References


96. Kretsinger, R. H. (1976) *Ann. Rev. Biochem.* **45**, 239–266
96a. Pawlowski, K., Bierzynski, A., and Godzik, A. (1996) *J. Mol. Biol.* **258**, 349–366
97. Ikura, M. (1996) *Trends Biochem. Sci.* **21**, 14–17
98. Kretsinger, R. H., and Nockolds, C. E. (1973) *J. Biol. Chem.* **248**, 3313–3326
99. McPhalen, C. A., Sielecki, A. R., Santarsiero, B. D., and James, M. N. G. (1994) *J. Mol. Biol.* **235**, 718–732
99a. Declercq, J.-P., Evrard, C., Lamzin, V., and Parello, J. (1999) *Protein Sci.* **8**, 2194 – 2204
99b. Biekofsky, R. R., Martin, S. R., Browne, J. P., Bayley, P. M., and Feeney, J. (1998) *Biochemistry* **37**, 7617 – 7629
100. Drake, S. K., Zimmer, M. A., Miller, C. L., and Falke, J. J. (1997) *Biochemistry* **36**, 9917–9926
101. Akke, M., Forsén, S., and Chazin, W. J. (1995) *J. Mol. Biol.* **252**, 102–121
102. Denisov, V. P., and Halle, B. (1995) *J. Am. Chem. Soc.* **117**, 8456–8465
103. Heizmann, C. W., and Hunziker, W. (1991) *Trends Biochem. Sci.* **16**, 98–103
104. Ikemoto, N., Ronjat, M., Mészáros, L. G., and Koshita, M. (1989) *Biochemistry* **28**, 6764–6771
105. Meldolesi, J., and Pozzan, T. (1998) *Trends Biochem. Sci.* **23**, 10–14
106. Randolph, J. T., McClure, K. F., and Danishefsky, S. J. (1995) *J. Am. Chem. Soc.* **117**, 5712–5719
107. Fliegel, L., Burns, K., MacLennan, D. H., Reithmeier, R. A. F., and Michalak, M. (1989) *J. Biol. Chem.* **264**, 21522–21528
108. Cohen, P., and Klee, C. B., eds. (1988) *Calmodulin*, Elsevier, Amsterdam
109. James, P., Vorherr, T., and Carafoli, E. (1995) *Trends Biochem. Sci.* **20**, 38–42
110. Lafitte, D., Capony, J. P., Grassy, G., Haiech, J., and Calas, B. (1995) *Biochemistry* **34**, 13825–13832
111. Chattopadhyaya, R., Meador, W. E., Means, A. R., and Quiocho, F. A. (1992) *J. Mol. Biol.* **228**, 1177–1192
112. Pedigo, S., and Shea, M. A. (1995) *Biochemistry* **34**, 10676–10689
113. da Silva, A. C. R., and Reinach, F. C. (1991) *Trends Biochem. Sci.* **16**, 53–57
114. Gagné, S. M., Li, M. X., and Sykes, B. D. (1997) *Biochemistry* **36**, 4386–4392
115. Trigo-Gonzalez, G., Awang, G., Racher, K., Neden, K., and Borgford, T. (1993) *Biochemistry* **32**, 9826–9831
116. Spyracopoulos, L., Li, M. X., Sia, S. K., Gagné, S. M., Chandra, M., Solaro, R. J., and Sykes, B. D. (1997) *Biochemistry* **36**, 12138–12146
117. Travé, G., Lacombe, P.-J., Pfuhl, M., Saraste, M., and Pastore, A. (1995) *EMBO J.* **14**, 4922–4931
118. Näthke, I., Hill, B. L., Parham, P., and Brodsky, F. M. (1990) *J. Biol. Chem.* **265**, 18621–18627
119. Pley, U. M., Hill, B. L., Alibert, C., Brodsky, F. M., and Parham, P. (1995) *J. Biol. Chem.* **270**, 2395–2402
120. Engel, J., Taylor, W., Paulsson, M., Sage, H., and Hogan, B. (1987) *Biochemistry* **26**, 6958–6965
121. Engel, E., Richter, K., Obermeyer, G., Briza, P., Kungl, A. J., Simon, B., Auer, M., Ebner, C., Rheinberger, H.-J., Breitenbach, M., and Ferreira, F. (1997) *J. Biol. Chem.* **272**, 28630–28637
122. Schäfer, B. W., and Heizmann, C. W. (1996) *Trends Biochem. Sci.* **21**, 134–140
123. Drohat, A. C., Baldisseri, D. M., Rustandi, R. R., and Weber, D. J. (1998) *Biochemistry* **37**, 2729 – 2740
123a. Gribenko, A. V., and Makhatadze, G. I. (1998) *J. Mol. Biol.* **283**, 679 – 694
123b. Brodersen, D. E., Nyborg, J., and Kjeldgaard, M. (1999) *Biochemistry* **38**, 1695–1704
124. Christiansen, W. T., Tulinsky, A., and Castellino, F. J. (1994) *Biochemistry* **33**, 14993–15000
125. Schwalbe, R. A., Ryan, J., Stern, D. M., Kisiel, W., Dahlbäck, B., and Nelsestuen, G. L. (1989) *J. Biol. Chem.* **264**, 20288–20296
126. Demange, P., Voges, D., Benz, J., Liemann, S., Göttig, P., Berendes, R., Burger, A., and Huber, R. (1994) *Trends Biochem. Sci.* **19**, 272–276
127. Saurel, O., Cézanne, L., Milon, A., Tocanne, J.-F., and Demange, P. (1998) *Biochemistry* **37**, 1403–1410
128. Moss, S. E., and Crumpton, M. J. (1990) *Trends Biochem. Sci.* **15**, 11–12
129. Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehmann, M. S., Grübel, G., Legrand, J.-F., Als-Nielsen, J., Colman, D. R., and Hendrickson, W. A. (1995) *Nature (London)* **374**, 327–336
130. Vyas, N. K., Vyas, M. N., and Quiocho, F. A. (1988) *Science* **242**, 1290–1295
131. Stuart, D. I., Acharya, K. R., Walker, N. P. C., Smith, S. G., Lewis, M., and Phillips, D. C. (1986) *Nature (London)* **324**, 84–87
132. Williams, R. J. P. (1991) *Trends Biochem. Sci.* **16**, 206

## Study Questions

1. a) From $\Delta G°$ for hydrolysis of sucrose (Table 6-6) calculate the equilibrium constant

$$K = [\text{glucose}][\text{fructose}] / [\text{sucrose}]$$

at 25°C. Call this hydrolysis reaction 1.
   b) Is the sucrose in a 1 M solution stable? Explain.
   c) If acid is added to a 1 M sucrose solution to catalyze its hydrolysis, what will be the final sucrose concentration at equilibrium? (Assume that concentrations equal activities for the purpose of these calculations.)
   d) Reaction 2 is the hydrolysis of α-D-glucose 1-phosphate to glucose and inorganic phosphate ($P_i$). Using $\Delta G°$ for this reaction (Table 6-6) calculate the equilibrium constant.
   e) Sucrose phosphorylase from the bacterium *Pseudomonas saccharophila* catalyzes the following reaction (reaction 3):

Sucrose + $P_1$ →
α-D-glucose 1-phosphate + fructose

Calculate the equilibrium constant and the standard Gibbs energy change at 25°C for reaction 3 from the equilibrium constants obtained above for reactions 1 and 2. Show that $\Delta G°$ for reaction 3 = $\Delta G°$ of reaction 1 - $\Delta G°$ of reaction 2.
   f) Could the bacterium carry out reaction 3 in the following two consecutive steps? Explain.

Sucrose → glucose + fructose
Glucose + $P_i$ → glucose 1-phosphate

2. For each of the following reactions, state whether the equilibrium constant will be between 0.1 and 10 (i.e., about one), greater than 100, or less that 0.01. Assume that the pH is constant at 7.0.
   a) $2\ ADP^{3-} \rightarrow ATP^{4-} + AMP^{2-}$
   b) $ATP^{3-}$ + glucose → glucose 6-phosphate$^{2-}$ + $ADP^{2-} + H^+$
   c) $ADP^{2-} + HPO_4^{2-} + H^+ \rightarrow ATP^{3-}$
   d) Glucose 6-phosphate$^{2-}$ → fructose 6-phosphate$^{2-}$
   e) Phosphoenolpyruvate$^{3-}$ + glucose → glucose 1-phosphate$^{2-}$ + pyruvate$^-$

3. The combustion of 1 mol of solid urea to liquid water and gaseous carbon dioxide and nitrogen ($N_2$) in a bomb colorimeter at 25°C (constant volume) liberated 666 kJ of heat energy. Calculate $\Delta H$, the change in heat content (enthalpy), for this reaction.

4. Using data of Table 6-4 calculate $\Delta G'$ (pH 7) for the following reactions:
   a) Glucose → 2 lactate$^-$ + 2 $H^+$
   b) $2\ NH_4^+ + HCO_3^- \rightarrow$ urea + 2 $H_2O$ + $H^+$
   c) 2-oxoglutarate$^{2-}$ + 1/2 $O_2$ + CoA + $H^+$ → succinyl-CoA + $H_2O$ + $CO_2$

5. What is the ionic strength of a 0.2 M solution of NaCl? of 0.2 M $Na_2SO_4$?

6. The [ATP] / [ADP] ratio in an actively respiring yeast cell is about 10. What would be the intracellular [3-phosphoglycerate] / [1,3-bisphosphoglycerate] ratio have to be to make the phosphoglycerate kinase reaction (Fig. 9-7, reaction 7) proceed toward 1,3-bisphosphoglycerate synthesis at 25°C, pH 7?

7. a) Using data from Table 6-8 determine the equilibrium constant for the reaction between malate and methylene blue, assuming all reactants present initially at the same concentration. Indicate clearly the direction of the reaction for which the Gibbs energy change is written.
   b) Calculate the percentage of the reduced (leuco) form of methylene blue present at pH 7 and 25°C in a system for which the measured electrode potential is 0.065 V.

8. $NAD^+$ is a coenzyme for both pyruvate dehydrogenase and ethanol dehydrogenase. Using the values of $E_o'$ from Table 6-8 calculate the Gibbs energy change and the equilibrium constant for the reaction.

Lactate$^-$ + acetaldehyde → pyruvate$^-$ + ethanol

9. Consider the oxidation of acetate at 25°C:

$$CH_3COO^- (0.1\ M) + 2\ O_2 (g, 0.2\ atm) + H^+ (10^{-7}\ M) \rightarrow 2\ H_2O (l) + 2\ CO_2 (g)$$

   a) What is the equilibrium pressure of $CO_2$ if the reaction is not coupled to any other reaction?
   b) What is the equilibrium pressure of $CO_2$ if the reaction is coupled to the formation of 0.01 M ATP from 0.02 M ADP and 0.01 M $HPO_4^{2-}$ in the citric acid cycle?
   c) What do the above calculations tell you about the prospects of gaining 100% efficiency of energy storage in ATP from the citric acid cycle?
   d) If the actual pressure of $CO_2$ is 0.01 atm, what is the efficiency of energy storage under the conditions in (b)?

## Study Questions

10. The equilibrium constant for the following reaction, which is catalyzed by creatine kinase, has been determined by chemical analysis. The data are given below. [S. A. Kuby and E. A. Noltman, in *The Enzymes*, 2nd ed. (P. D. Boyer, H. Lardy, and K. Myrbäck, eds), Vol. VI, pp. 515–602. Academic Press, New York, 1962]

$$ATP^{4-} + \text{creatine} \rightarrow ADP^{3-} + \text{creatine phosphate}^{2-} + H^+$$

| $t$ (°C) | $K$ |
|---|---|
| 20 | $6.30 \times 10^{-9}$ |
| 30 | $5.71 \times 10^{-9}$ |
| 38 | $5.47 \times 10^{-9}$ |

a) What are $\Delta G°$, $\Delta H°$, and $\Delta S°$ for the reaction at 25°C?
b) What are $\Delta G'$, $\Delta H'$, and $\Delta S'$ (pH 7) for the reaction at 25°C?
c) What are $\Delta G'$, $\Delta H'$, and $\Delta S'$ (pH 7) for the hydrolysis of creatine phosphate at 25°C?

11. The following reaction was carried out in a calorimeter at 25°C in 0.1 M phosphate buffer at pH 7.4 in the presence of a particulate suspension containing the mitochondrial electron transport system [M. Poe, H. Gutfreund, and R. W. Estabrook, *ABB* **122**, 204–211 (1967)]:

$$NADH + H^+ + 1/2\ O_2\ (\text{aq., sat.}) \rightarrow NAD^+ + H_2O$$

The oxygen consumption was monitored continuously with an oxygen electrode. The temperature was monitored simultaneously with a thermocouple immersed in the solution. At the start of the reaction 96 µmol NADH was added to 29.0 ml buffer containing $O_2$. A nearly zero-order reaction was observed with the rate of $O_2$ consumption of 6.87 µmol/min and the rate of temperature rise of 0.01171 K/min. The heat capacity of the calorimeter and contents was 254.6 J/K. What is $\Delta H$ for the above reaction? NOTE: The $H^+$ is supplied by the phosphate buffer, which has a $\Delta H$ of dissociation of 5.4 kJ $mol^{-1}$.

12. Enthalpy and Gibbs energy changes for the following reaction at 25°C are given in Table 6-6.

$$ATP^{4-}\ (1\ M) + H_2O \rightarrow ADP^{3-}\ (1\ M) + H^+\ (10^{-7}\ M) + HPO_4^{2-}\ (1\ M)$$

a) How much heat is evolved at constant temperature and pressure if the reaction takes place in a test tube without doing any work other than $p\,\Delta V$ work?
b) How much heat is evolved or absorbed by the foregoing reaction if it is coupled with 100% efficiency to an endergonic reaction?
c) What efficiency of coupling to an endergonic reaction is required in order that the foregoing reaction neither evolve nor absorb heat?

13. Microorganisms use a great variety of fermentation reactions for obtaining energy. Could the following reaction be used for such a purpose? Explain the reasons for your answer.

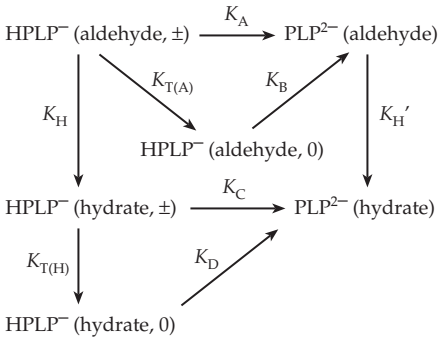
$$C_4H_6O_4\ (\text{succinic acid}) + H_2O \rightarrow C_3H_8O_3\ (\text{glycerol}) + CO_2$$

14. a) Calculate the work done in kJ and in kcal by a 70 kg person in climbing up stairs three stories (13 m).
b) Calculate how much ATP (in mmol and in grams) will be needed for the climb if muscles can use the ATP with 50% efficiency and if the phosphorylation state $R_p$ = [ATP] / [ADP] [$P_i$] is $10^3\ M^{-1}$. The standard value of $\Delta G^{°'}$ (pH 7) for hydrolysis of MgATP to MgADP and $P_i$ may be taken as −31 kJ/mol at 37°C.
c) How much of this ATP could be provided by transfer of phospho groups from the stored creatine phosphate (Cr-*P*) to ADP in muscle? Assume that [Cr-*P*] = 20 mM and may fall to 10 mM during the climb. $\Delta G^{°'}$ (pH 7) for hydrolysis of creatine phosphate is about −43 kJ/mol. If the creatine kinase reaction attains equilibrium what will the value of $R_p$ be for the adenylate system?

15. Acid–base titration gave the following three p$K_a$ values for cysteine: 1.70, 8.36, and 10.53. Spectrophotometric data allowed H. B. F. Dixon and K. F Tipton (1973, *Biochem J.* **133**, 837–842) to estimate the following ratio of tautomeric species at pH 9.4 [$^-OOC - C(NH_3^+) - CH_2S^-$] / [$^-OOC - C(NH_2) - CH_2SH$] = 2.12, as is also shown in Fig. 6-5. Verify and assign the four microscopic p$K_a$ values.

## Study Questions

16. The coenzyme pyridoxal phosphate (PLP; structure on p. 740) has p$K_a$ values of 3.62, 6.10, and 8.33, as determined by acid–base titration or by spectrophotometric titration. The p$K_a$ of 6.10 belongs primarily to the phosphate group and is nearly independent of the others. However, the p$K_a$ values of 3.62 and 8.33 are shared by the protonated ring nitrogen and the phenolic –OH group. PLP exists as an equilibrium mixture of aldehyde together with its covalent hydrate (see Eq. 13-1). The equilibrium constants for hydrate formation ($K_h$ = [hydrate] / [aldehyde]) are independent of pH but differ for each ionization state of the ring. Consider only the equilibria in the pH range 7–12.



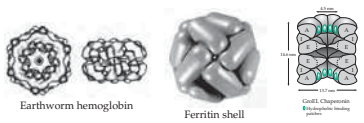

The UV-visible absorption spectrum of the monoprotonated form of PLP was divided mathematically into individual bands for the aldehyde with dipolar ionic ring (±), the aldehyde tautomer with an uncharged ring (0) and the hydrate of the dipolar ion. The following fractions were estimated (Harris *et al.*, 1976, *Biochem. Biophys. Acta.* **521**, 181–194.

| | |
|---|---|
| Aldehyde (±) | 52% |
| Aldehyde (0) | 28% |
| Hydrate (±) | 16% |
| Hydrate (0) | 4%, estimated indirectly |

From 300 MHz $^1$H NMR spectra areas for the following hydrogen atoms were estimated by Robitaille.

| | pH 7 | pH 12 |
|---|---|---|
| 4'-H aldehyde | .86 | .900 |
| 4'-H hydrate | .14 | .012 |
| 6-H, aldehyde | .96 | 1.059 |
| 6-H, hydrate | .17 | .014 |
| 2'-$CH_3$, aldehyde | 2.49 | 2.96 |
| 2'-$CH_3$, hydrate | .51 | .038 |

Evaluate the tautomerization constants $K_{T(A)}$ and $K_{T(H)}$, the hydration constants $K_H$ and $K_{H'}$, and the microscopic p$K_a$ values p$K_A$, p$K_B$, p$K_C$, and p$K_D$ in the foregoing scheme. Assume that the hydration ratios $K_h$ and $K_{h'}$ are the same in $H_2O$ (spectrophotometric data) and in $D_2O$ (NMR data).




Some ways in which protein subunits associate. (Left) The 3.66 MDa hemoglobin of the earthworm *Lumbricus terrestris* contains 144 globin subunits organized as 12 cylindrical disulfide-linked dodecamers. Two 6-dodecamer layers, each a ring with 6-fold cyclic symmetry, lie back-to-back. This reconstructed particle also contains three types of linker proteins in the center region. From Lamy *et al.* (2000) *J. Mol. Biol.* **298**, 638. (Center) The iron storage protein ferritin is formed from 24 19- to 21-kDa 4-helix-bundle subunits with cubic symmetry. As many as 4500 atoms of iron, as hydrated iron oxide, may be stored in the internal cavity. See Fig. 7-13. From Trikha *et al.* (1995) *J. Mol. Biol.* **248**, 954. Courtesy of Elizabeth Theil. (Right) The 2 MDa molecular chaperone GroEL consists of two back-to-back 7-subunit rings, each subunit formed from domains E, I and A. A 7-subunit cap of the smaller GroES may cover either end to form a compartment in which polypeptides fold. See Box 7-A.

# Contents

325 **A. Describing Binding Equilibria**
325 1. Analyzing Data
327 2. Multiple Binding Sites on a Single Molecule
328 *Microscopic binding constants and statistical effects*
329 *Electrostatic repulsion: anticooperativity*
330 3. Cooperative Processes
332 **B. Complementarity and the Packing of Macromolecules**
332 1. Rings and Helices
332 *Molecules with cyclic symmetry*
334 *Helical structures*
334 *Filamentous bacteriophages*
334 *A rod-shaped plant virus*
335 *Bacterial pili*
336 *The thin filaments of muscle*
337 2. Oligomers with Twofold (Dyad) Axes
337 *Paired interactions*
338 *Dihedral symmetry*
342 *Oligomers with cubic symmetry (polyhedra)*
344 *Asymmetry and quasi-equivalence in oligomers*
345 *Quasi-equivalence in virus coats*
348 *Regulatory subunits and multienzyme complexes*
349 **C. Cooperative Changes in Conformation**
349 1. Unequal Binding of Substrate and "Induced Fit"
350 2. Binding Equilibria for a Dimerizing Protein
350 *The Monod–Wyman–Changeux (MWC) model*
351 *The induced fit model*
352 *One conformational state dissociated*
352 3. Higher Oligomers
353 **D. The Oxygen-Carrying Proteins**
353 1. Myoglobin and Hemoglobin
353 *The binding of oxygen*
355 *Structural changes accompanying oxygen binding*
357 *The Bohr effect and allosteric regulators*
358 *Carbon monoxide, cyanide, and nitric oxide*
359 2. Abnormal Human Hemoglobins
362 3. Comparative Biochemistry of Hemoglobin
362 **E. Self-Assembly of Macromolecular Structures**
362 1. Bacteriophages
367 2. "Kringles" and Other Recognition Domains and Motifs
368 **F. The Cytoskeleton**
369 1. Intermediate Filaments
369 2. Microfilaments
370 3. Microtubules

373 **References**
377 **Study Questions**

**Boxes**

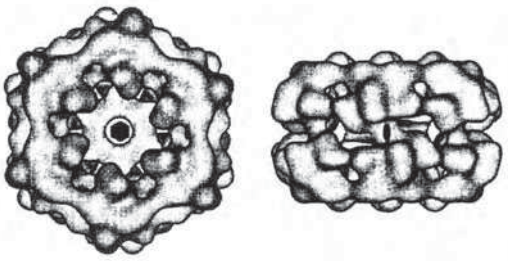
339 Box 7-A Life and Death for Proteins: Chaperonins and Proteasomes
360 Box 7-B Sickle Cell Disease, Malaria, and Blood Substitutes
363 Box 7-C The T-Even Bacteriophages
371 Box 7-D Mitosis, Tetraploid Plants, and Anticancer Drugs

**Tables**

329 Table 7-1 Binding Constants of Protons to Dianions of Dicarboxylic Acids
355 Table 7-2 Thermodynamic Functions for Oxygenation of Hemoglobin
367 Table 7-3 A Few Well-Known Structural Domains

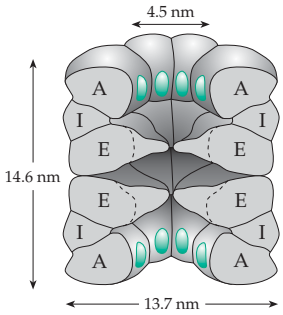
# How Macromolecules Associate

# 7



Earthworm hemoglobin

Ferritin shell




GroEL Chaperonin
Hydrophobic binding patches

The complicated shapes and internal structures of cells are determined to a large extent by the way in which proteins and other macromolecules are bonded one to another. In addition, intimate association of macromolecules is essential to such biological processes as the motion of flagella, the contraction of muscle, the action of antibodies, the transmission of nerve impulses, the replication of DNA, and the synthesis of proteins. Equally important is the binding of small molecules to large ones. In this chapter we will first examine methods of measuring binding with an emphasis on protons and small molecules. Then we will consider the ways in which macromolecules stick together as well as the role of conformational changes within macromolecules.

## A. Describing Binding Equilibria

In previous discussions of pH we have dealt with dissociation constants, but in this section we will use formation constants $K_f$, where $K_f = 1/K_d$. Measurement of the strength of association of molecules is an everyday aspect of modern biochemical research. It may be important to know how strongly a hormone binds to a receptor in a cell membrane or how well a feedback inhibitor binds to an enzyme to determine whether the interaction is significant physiologically. The binding of $O_2$ to hemoglobin and other oxygen carriers is vitally important, but the description of these oxygenation reactions is mathematically complex. This is especially so because we must consider effects of pH changes and of changing concentrations of allosteric effectors on the binding equilibria.

In considering such equilibria we must first examine the individual interactions of different domains of a protein, one with another. These can be described by association constants or, alternatively, by the Gibbs energy changes for the association reaction.[1,2] The average kinetic energy of motion of a molecule in solution is about $3/2k_BT$, where $k_BT$ is Boltzmann's constant. For one mole the kinetic energy is $3/2RT$ or 3.7 kJ (0.89 kcal) $mol^{-1}$ at 25°C. Thus, if $K_f = 10\ M^{-1}$ ($\Delta G° = -5.7$ kJ $mol^{-1}$ or $-1.36$ kcal $mol^{-1}$) the binding energy is only slightly in excess of the thermal energy of the molecules and the complex is weakly bound. In this instance, if X and P are both present in $10^{-4}$ molar concentrations (typical enough for biochemical systems), only 0.1% of the molecules will exist as the complex ([complex] = $K_f$[X][P]). If the formation constant is higher by a factor of 1000, i.e., $K_f = 10^4\ M^{-1}$ ($\Delta G° = -22.8$ kJ $mol^{-1}$), 38% of the molecules will exist as the complex; while if $K_f = 10^7\ M^{-1}$ (extremely strong binding, $\Delta G° = -40$ kJ or $-9.55$ kcal $mol^{-1}$), 97% of the molecules will be complexed.

### 1. Analyzing Data

The extent of binding of a molecule X to another molecule P (Eqs. 7-1, 7-2) is measured by varying the concentrations of X and P and observing changes in the concentration of the complex [PX]. The first

$$X + P \rightleftarrows PX \qquad (7\text{-}1)$$

$$K_f = [PX]/[P][X]) \qquad (7\text{-}2)$$

prerequisite is to find a measurable property that is different for the complex than for either of the free components. For example, the complex may be colored and the components colorless. More commonly, the complex simply has a different light absorbance (A) at

a certain wavelength than do the components. Likewise, the circular dichroism or the chemical shift of a peak in the NMR spectrum may change. If P is an enzyme, only the complex PX will undergo decomposition to products. Sometimes (but not always) the rate of breakdown of PX (the enzyme–substrate complex) to form products will be relatively slow compared to the rate at which the equilibrium between X, P, and PX is established. In this case the concentration of complex PX will be proportional to the observed rate of formation of product.

Whatever change of property is measured, its value will increase with increasing concentrations of X if the total concentration of the macromolecule P is kept constant. In the usual experimental design, the molar concentration of P is small and it is possible to increase the concentration of X to quite large values. When this is done, it is usually observed that at high enough values of [X] almost all of the P is converted to PX, and the change being measured (e.g., $\Delta A$ for increased light absorption) no longer increases. This effect is known as **saturation** and is observed in most binding studies and also in many physiological phenomena.

The property being measured ($\Delta A$) reaches a maximum value $\Delta A_{max}$ at saturation and when all of compound P has been converted to PX. The ratio of [PX] to the total concentration of all forms of P present $[P]_t$ is known as the **saturation fraction** and is often given the symbol Y. If P has more than one binding site for X, Y is defined as the fraction of the total binding sites occupied. If $n$ is the number of sites per molecule, the total number of sites is $n[P]$. The value of Y is often taken as $\Delta A/\Delta A_{max}$, an equality that holds for multisite macromolecules only if the change in $A$ is the same for each successive molecule of X added. This is not always true, but when it is Eq. 7-3 is followed.

$$\sum_i i\,\frac{[PX_i]}{n[P]_t} = Y = \frac{\Delta A}{\Delta A_{max}} \qquad (7\text{-}3)$$

Here $i$ represents the number of ligands X bound to P and may vary from 0 to $n$. When $n = 1$ the saturation fraction Y and $\Delta A$ are related to the concentration of free unbound X and the formation constant as follows:

$$Y = \frac{K_f[X]}{1 + K_f[X]} \qquad \Delta A = \frac{\Delta A_{max}\, K_f[X]}{1 + K_f[X]} \qquad (7\text{-}4)$$

A plot of Y or $\Delta A$ against [X] is shown in Fig. 7-1. This kind of plot is sometimes called an **adsorption isotherm** because it describes binding only at a constant temperature. Notice, from both Fig. 7-1 and Eq. 7-4, that Y reaches a value of 0.5 when [X] is just equal to $1/K_f$ (or to $K_d$). Note also that as [X] increases saturation is reached slowly and that even at the point representing the highest concentration of X ($8/K_f$ in Fig. 7-1) saturation is less than 90%. Since in the usual experimental situation, we do not know Y but only $\Delta A$, it is difficult to estimate the limiting value $\Delta A_{max}$ from a plot of this type unless $K_f$ is very high. However, we need to know $\Delta A_{max}$ to evaluate $K_f$. For this reason, plots like that of Fig. 7-1 are seldom used, this one being included mainly to illustrate a point of nomenclature. The curve shown in Fig. 7-1 is a rectangular hyperbola, and the type of saturation curve shown is frequently referred to as **hyperbolic**. This is in contrast to certain other binding curves (Section 3) which, when plotted in this way, are **sigmoidal** (S-shaped).

A better type of plot is often that of Y against log [X] (Fig. 7-2). It has the following features. (1) The curve is symmetric about the midpoint at log [X] = log $K_f$. (2) No matter how high or low the concentration range used in the experiments, it is easy to choose a scale that puts all the points on the same sheet of paper. (3) Spacing between points tends to be more uniform than in a plot against [X]; e.g., compare Figs. 7-1 and 7-2 for which the experimental points represent the same data and for which values of [X] for successive points are each twofold greater than the preceding one. (4) The same logarithmic scale can be used for all compounds, no matter how strong or weak the binding, and the same shape curve is obtained for all 1:1 complexes. The midpoint slope, $dY/d\log[X]$, is 0.576; the change in log [X] in going from 10 to 90% saturation is 1.81. The curve is familiar to most chemists because it is frequently used for pH titration curves in which pH substitutes for –log [X]. To represent a complex with tighter binding, the curve is simply moved to the left, and for weaker binding, it is moved to the right.

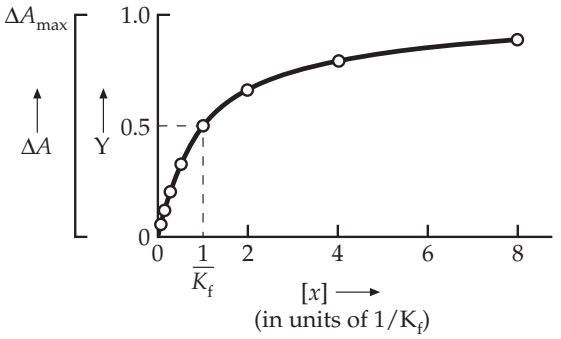



**Figure 7-1** An adsorption isotherm, a plot of the saturation fraction Y or of some change in a measured property $\Delta A$ vs [X], the concentration of a substance that binds reversibly to a macromolecule. The curve is hyperbolic and $[X] = 1/K_f$ when Y = 0.5.

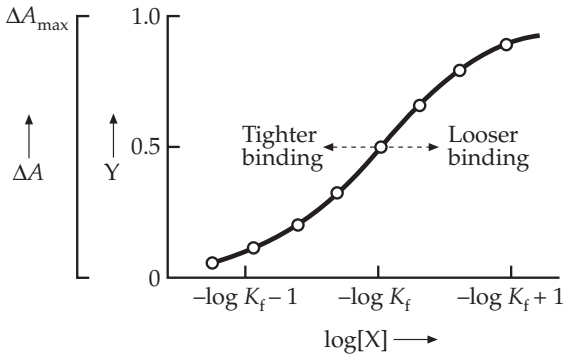



**Figure 7-2** A saturation curve plotted on a logarithmic scale for [X]. The data points are the same as those used in Fig. 7-1.

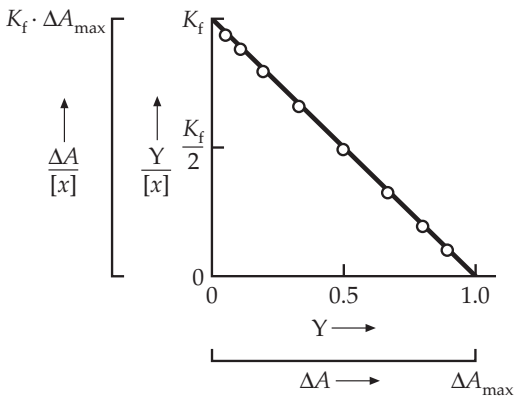



**Figure 7-3** A Scatchard plot of the same data shown in Figs. 7-1 and 7-2. This is the best of the linear plots for studying binding.

Saturation data are often plotted in yet another form known as the **Scatchard plot** (Fig. 7-3). The value of $\Delta A/[X]$ (or of $Y/[X]$) is plotted against $\Delta A$ (or Y) and a straight line is fitted to the points, preferably using the "method of least squares." The intercept on the $x$ axis and the slope of the fitted line give values of $\Delta A_{max}/K_f$ and $K_f$, respectively, as indicated by Eq. 7-5, which follows directly from Eq. 7-4.

$$\begin{aligned} Y/[X] &= K_f - YK_f \\ \Delta A/[X] &= \Delta A_{max} K_f - \Delta A K_f \end{aligned} \qquad (7\text{-}5)$$

The Scatchard plot is the best of the various linear transformations of the saturation equation and is preferred to "double reciprocal plots" analogous to that shown in Fig. 9-3.

Scatchard's original equation was formulated to deal with the binding of two or more ligands to a single macromolecule.[3–5] If we let $[X]_b$ represent the concentration of bound X and $[P]_t$ the total molar concentration of the protein or other macromolecule and if there is only one binding site on the protein, $[X]_b/[P]_t$ will equal Y. However, if there are $n$ independent binding sites that have the same binding constant $K_f$ Eq. 7-5a will hold.

$$\frac{[X]_b}{[P]_t} \cdot \frac{1}{[X]} = nK_f - \frac{[X]_b}{[P]_t} K_f \qquad (7\text{-}5a)$$

If $[X]_b/[P]_t[X]$ is plotted against $[X]_b/[P]_t$ the resulting linear plot will have an intercept of $K_f$ on the $y$ axis and $n$ on the $x$ axis. Thus, $n$ is directly apparent, whereas in Eq. 7-5 it is incorporated into Y. A problem arises if, as discussed in the next section, the multiple binding sites are not independent but interact. Curved Scatchard plots result and attempts to extract more than one binding constant can lead to very large errors. Before measuring saturation curves, the student should read additional articles or books on the subject.[2,6–10]

## 2. Multiple Binding Sites on a Single Molecule

A macromolecule may often be able to bind several molecules of a second compound X. Consider the case in which the macromolecule P binds successively one molecule of X, then a second, and a third, up to a total of $n$. We define the stepwise formation constants, $K_1, K_2, \ldots, K_n$, as follows:

$$\begin{aligned} P + X &\xrightarrow{K_1} PX \\ PX + X &\xrightarrow{K_2} PX_2 \ldots \\ PX_{n-1} + X &\xrightarrow{K_n} PX_n \end{aligned} \qquad (7\text{-}6)$$

Remember that these are reversible reactions even though unidirectional arrows are used. The general expression for the $i$th stepwise formation constant is given by Eq. 7-7.

$$K_i = \frac{[PX_i]}{[PX_{i-1}][X]} \qquad (7\text{-}7)$$

Remember that Y is the fraction of total binding sites saturated. The number of moles of X bound per mole of P is $n$Y and is obtained by summing the concentra-

tions [PX] + 2[$PX_2$] + . . . and dividing the sum of all the forms of P:

For two binding sites ($n = 2$)

$$2\bar{Y} = \frac{[PX] + 2[PX_2]}{[P] + [PX] + [PX_2]} \qquad (7\text{-}8)$$

For the general case

$$n\bar{Y} = \left( \sum_{i=1}^{n} i[PX_i] \right) \Big/ \left( [P] + \sum_{i=1}^{n} [PX_i] \right) \qquad (7\text{-}9)$$

The summations are over all of the integral values of $i$ from 1 to $n$. Now, by expressing each concentration, [$PX_i$], in terms of the concentrations [X] and [P] of *free* X and P, together with the stepwise formation constants, we obtain Eq. 7-10.

For $n = 2$

$$2\bar{Y} = \frac{K_1[X] + 2K_1K_2[X]^2}{1 + K_1[X] + K_1K_2[X]^2} \qquad (7\text{-}10)$$

A similar equation can be written for the general case. Note that the concentration of P does not appear in Eq. 7-10 and that Y is a function only of [X] and the stepwise formation constants. Such equations define the isotherms for binding of two or more molecules of X to P. From an experimental plot of Y (or of $\Delta A$) vs [X] or log [X], it is possible in favorable cases to determine the stepwise constants $K_1$, $K_2$, ..., $K_n$. However, this becomes quite complicated. To simplify Eq. 7-10 and the corresponding equation for the general case, we can group the constants together and designate the products of constants ($K_1$, $K_1K_2$, $K_1K_2K_3$, etc.) as $\psi_1$, $\psi_2$, . . ., $\psi_n$. Our equations are now as follows:

For $n = 2$

$$2\bar{Y} = \frac{\psi_1[X] + 2\psi_2[X]^2}{1 + \psi_1[X] + \psi_2[X]^2} \qquad (7\text{-}11)$$

For the general case

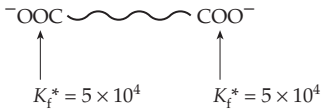
$$n\bar{Y} = \sum_{i=1}^{n} i\psi_i[X]^i \Big/ \left( 1 + \sum_{i=1}^{n} \psi_i[X]^i \right) \qquad (7\text{-}12)$$

From experimental data, it is usually easiest to first determine the $\psi$'s (there are $n$ of them), and then to calculate from the $\psi$'s the stepwise constants. For example:

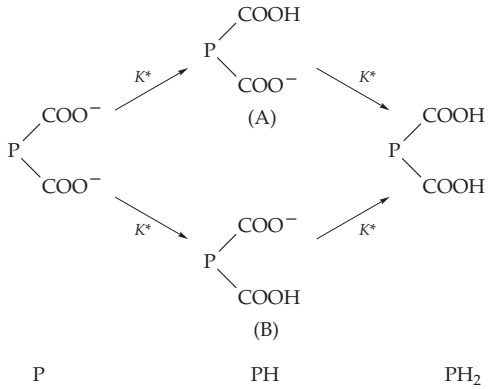
$$K_1 = \psi_1 \qquad K_2 = \psi_2/K_1 \text{, etc.} \qquad (7\text{-}13)$$

While Eq. 7-12, known as the **Adair equation**,[11] might seem to provide a complete description of the binding process, it usually does not. In many cases, there is more than one kind of binding site on a macromolecule and Eq. 7-12 tells us nothing about the distribution of the ligand X among different sites in complex PX. To consider this problem we must examine the *microscopic binding constants*.

***Microscopic binding constants and statistical effects.*** As discussed in Chapter 6, Section E,2, microscopic binding constants represent the constants for binding to specific individual sites. Now, consider a straight-chain dicarboxylic acid which has two *identical* binding sites for protons. If the chain connecting the two carboxylate anions is long enough, the carboxylate groups will be far enough apart that they do not influence each other through electrostatic interaction.




Each group will have a microscopic binding constant ($K_f^*$) of 5 x $10^4$ $M^{-1}$. The constant $K_f^*$ can also be called an **intrinsic binding constant**, because it is characteristic of a carboxylate group that is free of interactions with other groups. Intuition tells us (correctly) that, in its binding of protons, a solution of this dicarboxylic acid dianion will behave exactly like a solution of the monovalent anion R–$COO^-$ at twice the concentration. A single intrinsic binding constant suffices to describe both binding sites. It may seem surprising then that the **stepwise formation constants** (also called **stoichiometric** or **macroscopic** formation constants) $K_1$ and $K_2$ differ: $K_1$ = 10 x $10^4$ $M^{-1}$ and $K_2$ = 2.5 x $10^4$ $M^{-1}$. This fact reflects the so-called statistical effect. Either of the two carboxylate groups in the molecule can bind a proton in the first step to give two indistinguishable molecules, PH:




(7-14)

If we label the two forms of PH as A and B (Eq. 7-14) and consider that each one of them is independently in equilibrium with P through formation constant $K_f$*, we obtain Eq. 7-15 (which may be compared with Eqs. 6-75 and 6-76, which are written for *dissociation* constants).

$$K_1 = \frac{[PH]_A + [PH]_B}{[P][H^+]} = 2K^* \qquad \text{and} \qquad K_2 = K^*/2 \tag{7-15}$$

This result is related to probability and arises for the same reason that if you reach into a barrel containing 50% white balls and 50% black balls, you will pull out one of each just twice as often as you will pull out a pair of white or a pair of black. In the general case of $n$ equivalent binding sites, the microscopic formation constants $K_i$* are related[12,13] to the stepwise constants $K_i$ as follows:

$$K_i = \frac{(n + 1 - i)}{i} K_i^* \tag{7-16}$$

It is also easy to show,[14] using Eqs. 7-12 and 7-16 that for *n completely equivalent and independent* binding sites Eqs. 7-17 and 7-18 hold:

$$Y = \frac{K^*[X](1 + K^*[X])^{n-1}}{(1 + K^*[X])^n} \tag{7-17}$$

or

$$Y = \frac{K^*[X]}{1 + K^*[X]} \tag{7-18}$$

In this case the microscopic association constants are all identical and represent a single **intrinsic** constant applicable to all of the sites. In fact, Eq. 7-18 is identical to that for association of a single proton (or other ligand) with a single binding site, satisfying our intuitive notion that a set of $n$ completely independent binding sites should behave just like a solution of an $n$-fold more concentrated compound with a single binding site. Thus, our arithmetic has led us to a conclusion that was already obvious. However, it is rarely true that binding sites on a single macromolecule are completely independent; there is almost always *interaction* between them, and the equations that we have derived for evaluation of stepwise and intrinsic constants cannot be applied without modification.

***Electrostatic repulsion: anticooperativity.*** As we have seen, a hypothetical acid with an infinite distance between the carboxylate groups and log $K_f$* = 4.8 would have two macroscopic binding constants separated by the statistical distance (log 4 = 0.6). Compare these values with the observed binding constants for protons with the dianions of acids containing 7, 4, 2, and 1 $CH_2$ groups given in Table 7-1. For the longest chain (that of azelaic acid) the log $K_f$ values are not very different from those of the hypothetical long-chain acid. However, as the groups come closer together, the first binding constant is increased markedly because of the additional electrostatic attraction and the second is decreased. The spread between the two log $K_f$ values increases from 0.6 to as much as 2.9 as a result of interaction between the binding sites.

**TABLE 7-1**
**Binding Constants of Protons to Dianions of Dicarboxylic Acids[a]**

| Acid dianion | No. of $CH_2$ groups | log $K_1$ (p$K_2$) | log $K_2$ (association) (p$K_1$) (dissociation) |
|---|---|---|---|
| Hypothetical dianion with log $K^*$ = 4.8 | ∞ | 5.1 | 4.5 |
| Azelaic | 7 | 5.41 | 4.55 |
| Adipic | 4 | 5.41 | 4.42 |
| Succinic | 2 | 5.48 | 4.19 |
| Malonic | 1 | 5.69 | 2.83 |

[a] From R. P. Bell, (1973) *The Proton in Chemistry*, 2nd ed., p. 96. Cornell Univ. Press, Ithaca, New York

In malonic and succinic acids the first proton bound can be shared by both carboxyl groups through formation of a hydrogen bond. (See discussion in Chapter 9, Section D.) Additional factors operate in oxalic acid where the carboxyl groups are connected directly and for which p$K_a$ values (p$K_a$ = log $K_f$) are 4.19 and 1.23. In all of these examples the binding of the first proton makes it harder to bind a second proton. Such negative interaction or **anticooperativity** between binding sites is very common and always leads to a spread of the formation constants and a broadening of the curve of Y vs log [X]. This is shown graphically on the right side of Fig. 7-4 where the binding curves for protons with acetate ion and with succinic acid dianion are compared. Notice that binding of protons *increases* as log [$H^+$] increases, giving the curves an unfamiliar appearance when compared with the more familiar curve of *dissociation vs* pH.

Can we predict the p$K_a$ values in Table 7-1? With an appropriate dielectric constant chosen Eq. 2-8 can be applied. The difference between the two successive log $K_f$ values reduced by 0.6 (the statistical factor) is a

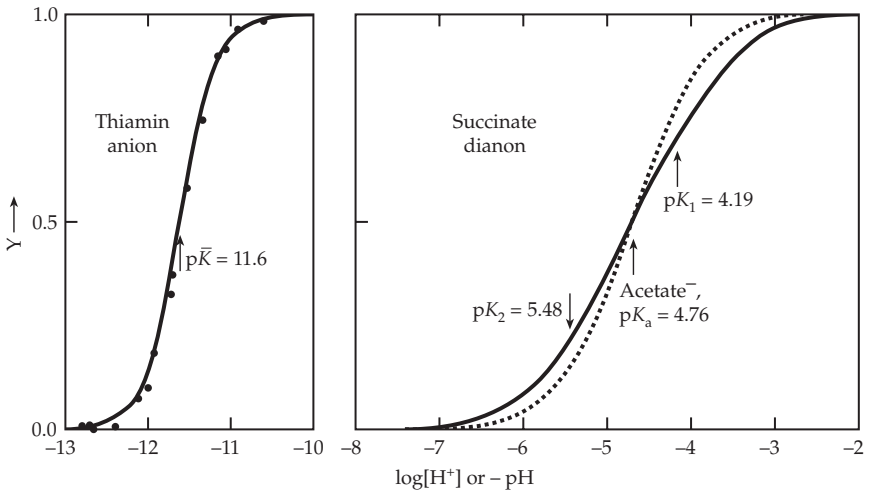
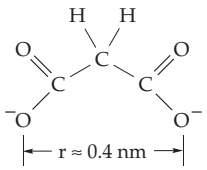



**Figure 7-4** Binding of protons to the thiamin anion, the succinate dianion, and the acetate anion. Acetate (dashed line) binds a single proton with a normal width binding curve. Succinate dianion binds two protons with anticooperativity, hence a broadening of the curve. The thiamin anion (yellow form, see Eq. 7-19) binds two protons with complete cooperativity and a steep binding curve.

measure of the electrostatic effect. For malonic acid this $\Delta pK_a$ is 2.25 and for succinic acid it is 0.69 (Table 7-1). In 1923, N. Bjerrum proposed that the value of $\Delta pK_a$ could be equated directly with the work needed to bring the two negative charges together to a distance representing the charge separation in the malonate dianion.




Thus, applying Eq. 6-31, $\Delta G = 5.708\ \Delta pK_a$ kJ mol$^{-1}$ = 12.84 kJ mol$^{-1}$ for malonate. Equating this with W in Eq. 2-8 and assuming a dielectric constant of 78.5 (that of water), the distance of charge separation $r$ is calculated to be 0.138 nm. This is much too small. The computation was improved by Westheimer and Kirkwood, who assumed a dielectric constant of 2.0 *within* the molecule. By approximating the molecule as an ellipsoid of revolution, they were able to make reasonably accurate calculations of electrostatic effects on $pK_a$ values.[15] Thus, for malonic acid Westheimer and Shookhoff[16] predicted $r$ = 0.41 nm for malonic acid dianion. Recently more sophisticated calculations[17] have been used to predict $pK_a$ values for the compounds in Table 7-1 and others.[18]

Electrostatic theory has also been used successfully to interpret titration curves of proteins in which the net negative or positive charge distributed over the surface of the protein varies continuously from high pH to low as more protons are added.[19]

Electrostatic effects can be transmitted extremely effectively through aromatic ring systems, a fact that explains some of the significance of heterocyclic aromatic systems in biochemical molecules. Consider the microscopic binding constant of the phenolate anion of pyridoxine as influenced by the state of protonation of the ring nitrogen. These are shown in Eq. 6-75, where $pK_a^* = 4.94$ and $pK_d^* = 8.20$ define the binding constants for protonation of a phenolate ion when the ring nitrogen is protonated or unprotonated, respectively. We see that $\Delta pK_a = 3.26$, even greater than that of malonic acid.

## 3. Cooperative Processes

Can it ever happen that interaction between groups leads to a *decrease* from the statistical separation between values of the stepwise constants instead of to an increase? At first glance, the answer seems to be no. A decreased separation would imply that the intrinsic binding constant for the second proton bound is higher than that for the first, but common sense tells us that the first proton ought to bind at the site with the highest binding constant. However, look at the experimental binding curve of protons with the anion of thiamin shown in Fig. 7-4. Instead of being broadened from the curve of acetate, it is just half as wide. The explanation depends upon some rather amusing chemistry of thiamine. Under suitable conditions, this vitamin can be crystallized as a yellow sodium salt, the structure of whose anion is shown in Eq. 7-19. Weak binding of a proton to one of the nitrogens as shown in Eq. 7-19 creates an electron deficiency at the adjacent carbon and the –S$^-$ anion adds to the C=N group, closing the ring to an unstable tricyclic form of thiamin. This tricyclic form can be observed in methanol and can be crystallized. It is unstable in water because the central ring can open, with the electrons flowing as indicated by the small arrows to create a strongly basic site on the same nitrogen. A second proton combines at this basic nitrogen with a high binding constant to form a cation.

*The key to the reversed order of strength of the binding*

*constants lies in the molecular rearrangements intervening between the two binding steps.*[20,21] In this particular case, we cannot measure the successive binding constants $K_1$ and $K_2$ directly from the titration curve because $K_2$ is almost two orders of magnitude larger than $K_1$. Consequently, the binding curve shown in Fig. 7-4 is (within the experimental error) twice as steep at the center (the slope is 2 x 0.576) as that for acetate ion and is accurately represented in Eq. 7-20. Comparison of Eq. 7-20 with Eq. 7-10 shows how the latter has been simplified because no significant concentration of the form PX is present in the cooperative case.

$$Y = \frac{\bar{K}^2[X]^2}{1 + \bar{K}^2[X]^2} \qquad (7\text{-}20)$$

This is only part of the story about the acid–base chemistry of thiamin. For the rest, see Chapter 14, Section D,1.

The binding of protons by the thiamin anion is an example of a **cooperative process**, so named because binding of the first proton makes binding of the second easier. Although relatively rare among small molecules, cooperative processes are very common and important in biochemistry.[22,23] A cooperative binding curve is sometimes referred to as **sigmoidal** because the plot of Y against [X] (the binding isotherm) is S shaped. The maximum possible cooperativity is observed when the binding of the first ligand enhances the affinity of all other sites so much that no species other than P and $PX_n$ are present in significant concentration.

It is easy to show that, for $n$ binding sites with such *completely cooperative binding*, the saturation fraction is:

$$Y = \frac{\bar{K}^n[X]^n}{1 + \bar{K}^n[X]^n} \qquad (7\text{-}21)$$

where $\bar{K} = (K_1 \ldots K_n)^{1/n}$. The midpoint slope in the binding curve (Y vs log [X]) is $0.576n$ and the change, $\Delta$log [X], between Y = 0.1 and Y = 0.9 is $1.81/n$.

Equation 7-21 can be rewritten as

$$Y/(1 - Y) = \bar{K}^n[X]^n \qquad (7\text{-}22)$$

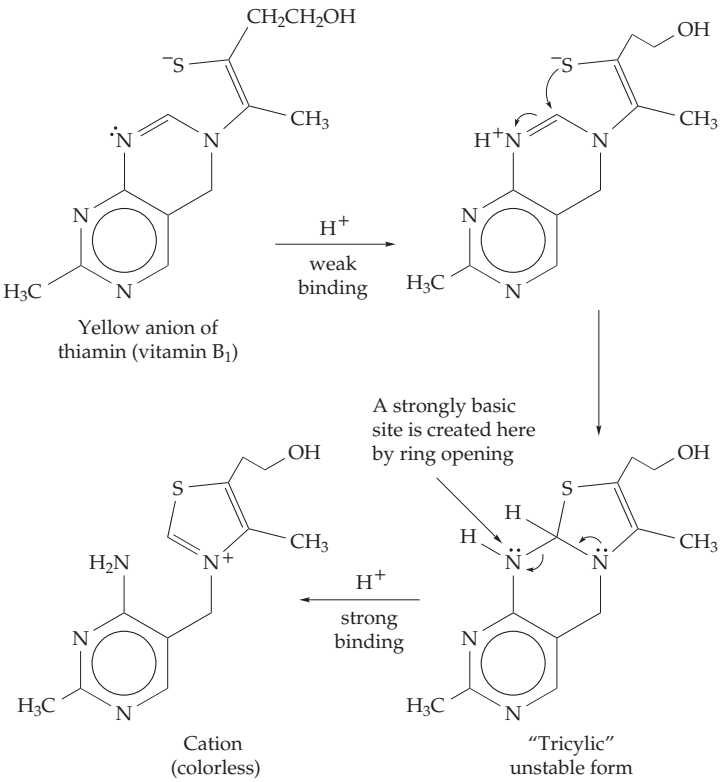
Taking logarithms (Eq. 7-22)

$$\log [Y/(1 - Y)] = n \log \bar{K} + n \log [X] \qquad (7\text{-}23)$$

A plot of log [Y/ (1 – Y)] vs log [X] is known as a **Hill plot**. According to Eq. 7-22, it is linear with a slope of $n$. Remember that this equation was derived for an ideal case of completely cooperative binding at $n$ sites.[24] However, Hill plots are often used to plot experimental data for systems in which cooperativity is incomplete. Thus, the experimentally measured slope of a Hill plot ($n_{Hill}$) is not an integer and is usually less than $n$, the number of binding sites. A comparison of $n_{Hill}$ with $n$ is often used as a measure of the degree of cooperativity: $n_{Hill}/n = 1.00$ for complete cooperativity but is less than one if cooperativity, is incomplete. An example of a very high degree of cooperativity is provided by the hexameric enzyme glutamate dehydrogenase, whose saturation curve for substrate displays $n_{Hill}$ approaching six.[25] It is not necessary to make a Hill plot to get $n_{Hill}$. From the usual binding curve of Y (or $\Delta A$) vs log [X] the midpoint slope can be measured with satisfactory precision. Alternatively, the difference, $\Delta$log [X], between 0.1 and 0.9 saturation can be evaluated and $n_{Hill}$ calculated from Eq. 7-24.

$$n_{Hill} = \frac{\text{midpoint slope}}{0.576} = \frac{1.81}{\Delta \log[X]} \qquad (7\text{-}24)$$

Binding curves sometimes show more than one step; in such cases Hill plots are not linear and no simple measure of cooperativity can be defined.




(7-19)

A second example of cooperativity is provided by the reversible denaturation of coiled peptide chains. Some proteins can be brought to a pH of 4 by addition of acid but without protonation of buried groups with intrinsic p$K_a$ values greater than four. When a little more acid is added, some less basic group is protonated, permitting the protein to unfold and to expose the more basic hidden groups. Thus, cooperative proton binding is observed. As in the case of thiamin the cooperativity depends upon the occurrence of a conformational change in the molecule linked to protonation of a particular group.

The reversible transformation between an α helix and a random coil conformation is also cooperative. In this case, once a helix is started, additional turns form rapidly and the molecule is completely converted into the helix. Likewise, once it unfolds it tends to unfold completely. Melting of DNA (Chapter 5) or, indeed, of any crystal is cooperative.[8] The stacking of nucleotides alongside a template polynucleotide can also be cooperative. For example, the binding of an adenylate residue to two strands of polyuridylic acid leads to cooperative formation of a triple-helical complex (Chapter 5, Section C,4). Here the stacking interactions make helix growth energetically easier than initiation of new helical regions.[26]

## B. Complementarity and the Packing of Macromolecules

Because the forces acting between them are weak, two molecules will cling together tightly only if there is a close fit between their surfaces. For a firm bond to be formed many atoms must be in contact and the two molecular surfaces must be *complementary* one to the other. If a "knob," such as a $-CH_3$ group, is present on one surface, there must be an appropriate hollow in the complementary surface. A positive charge in one surface is likely to be opposite a negative charge in the other. A proton donor group can form a hydrogen bond only if it is opposite a group with unshared electrons; nonpolar (hydrophobic) groups must be opposite each other if hydrophobic interaction is to occur. An important principle is that *two molecules with complementary surfaces tend to join together and interact, whereas molecules without complementary surfaces do not interact*. Watson called this "selective stickiness."[27] Selective stickiness permits the **self-assembly** of biological macromolecules having surfaces of complementary shape into fibers, tubes, membranes, and polyhedra. It also provides the means for specific pairing of purine and pyrimidine bases during the replication of DNA and during the synthesis of RNA and of proteins.

Complementarity of surfaces is equally important to the chemical reactions of cells. Each of these reactions is catalyzed by an enzyme, which contains reactive chemical groupings in the right places and in the right orientations to interact with and promote a chemical change in another molecule, the **substrate**. Specific catalysis is one of the most basic characteristics of living things. Enzymatic catalysis provides the basis not only for the reactions of metabolism but also for the movement of muscle fibers, the flowing of the cytoplasm in the ameba, and virtually all other biological responses. To understand these phenomena requires an examination of the structures of the macromolecules involved and of the ways in which they can fit together.

Just as the amino acids, sugars, and nucleotides are the building blocks for formation of proteins, polysaccharides, and nucleic acids, these three kinds of macromolecule are the units from which larger subcellular structures are assembled. Fibers, microtubules, virus "coats," and small symmetric groups of **subunits** in **oligomeric proteins** all result from the packing of macromolecules in well-defined ways, something that is often called **quaternary structure**.

### 1. Rings and Helices

Consider first the aggregation of identical protein subunits. While many protein molecules are nearly spherical, they are nevertheless asymmetric. In the drawings that follow the asymmetry is exaggerated, but the principles illustrated are valid. One easily observed lesson from nature is that even though living things are made up of asymmetric materials, a great deal of symmetry is evident.[28] At the molecular level the symmetry of crystalline arrays of atoms or molecules is described mathematically by the elements of symmetry present in **space groups** (p. 133). There are 230 of these but only 65 accommodate asymmetric objects (Chapter 3).[29] Two of the natural ways for identical asymmetric subunits to interact lead to rings and helices, respectively.

***Molecules with cyclic symmetry.*** Consider a subunit (**protomer**) of the shape shown in Eq. 7-25 and containing a region *a* that is complementary to the surface *j* on another part of the same molecule. Two such protomers will tend to stick together to form a dimer, region *a* of one protomer sticking to region *j* of the other. The dimer will still contain a free region *a* at one end and a region *j* at the other which are not involved in bonding. Other protomers can stick to these free ends. In some instances long chains can be formed. However, if the geometry is just right, a third subunit can fit in to form a closed ring (a trimer). Depending on the geometry of the subunits, the ring can be even smaller (a dimer) or it can be larger (a tetramer, pentamer, etc.). The bonding involved is

between two different regions (*a* and *j*) of a subunit and is sometimes described as **heterologous**.[30] To obtain a closed ring of subunits, the angle between the bonding groups *a* and *j* must be correct or the ring cannot be completed.

A ring formed using exclusively heterologous interactions possesses **cyclic symmetry**. The trimer in Eq. 7-25 has a **threefold axis**: Each subunit can be superimposed on the next by rotation through 360° / 3. The oligomer is said to have $C_3$ symmetry. Many real proteins, including all of those with 3, 5, or another uneven number of identical protomers, appear to be formed of subunits arranged with cyclic symmetry. An example is the cholera toxin from *Vibrio cholerae*, which forms a pentamer with an outer ring of subunits with $C_5$ symmetry (Fig. 7-5).

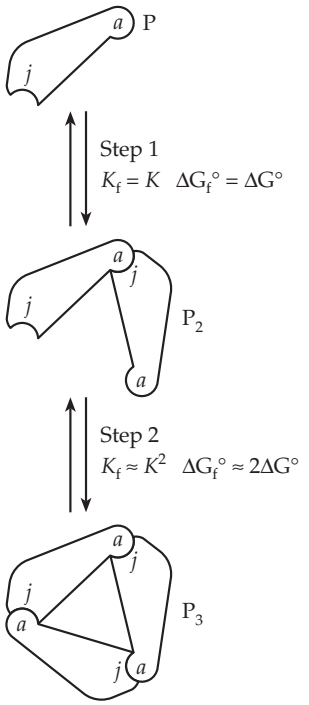
Now consider the quantitative aspects of heterologous interactions with ring formation. Let $K_f$ be the formation constant and $\Delta G^\circ$ the Gibbs energy change for the reaction of the *j* end of protomer P with the *a* end of a second protomer to form the dimer $P_2$ (Eq. 7-25).

In the second step (Eq. 7-25) a third protomer combines. It forms *two* new *aj* interactions. If we assume for this step that $\Delta G_f^\circ$ is 2 $\Delta G^\circ$, $K_f$ will be equal to $K^2$. The overall association constant for formation of a trimer from three protomers will be given by Eq. 7-26.

$$\begin{gathered} 3\ P \rightarrow P_3 \\ K_f = K^3 \\ \Delta G_f^\circ = 3\ \Delta G^\circ \end{gathered} \qquad (7\text{-}26)$$

This will be true only if $\Delta G^\circ$ for formation of both new *aj* bonds in the trimers is exactly the same as that for formation of the *aj* bond in the dimer. The reader may wish to criticize this assumption[30a,b,c] and to suggest conditions that might lead to overestimation or underestimation of $K_f$ for the trimer as calculated previously.

Now consider a hypothetical example: Protomer P is continuously synthesized by a cell. At the same time some subunits are degraded to a nonaggregating form via a second metabolic reaction. The two reactions are balanced so that [P] is always present at a steady state




(7-25)

value of $10^{-5}$ M. Suppose that a value for a single *aj* interaction of $K = 10^4$ (and $\Delta G^\circ = -22.8$ kJ mol$^{-1}$) governs aggregation to form dimers and trimeric rings. What concentration of dimers and of trimer rings will be present in the cell in equilibrium with the $10^{-5}$ concentration of P. Using Eq. 7-25 we see that the concentration of dimers $[P_2]$ is $10^4 \times (10^{-5})^2 = 10^{-6}$ M. (Note that the amount of material in this concentration of dimer is equivalent to $2 \times 10^{-6}$ M of the monomer units.) The concentration of rings $[P_3]$ is $(10^4)^3 \times (10^{-5})^3 = 10^{-3}$ M, equivalent to $3 \times 10^{-3}$ M of the monomer units. Thus, of the *total* P present in the cell ($10^{-5} + 0.2 \times 10^{-5} + 300 \times 10^{-5}$ M), 99.6% is associated with trimers, 0.33% is still monomers, and only 0.07% exists as dimers. Thus, the formation of two heterologous bonds simultaneously to complete a ring imparts a high degree of cooperativity to the association reaction of Eq. 7-25. We will find in a cell mostly either rings or monomer but little dimer.

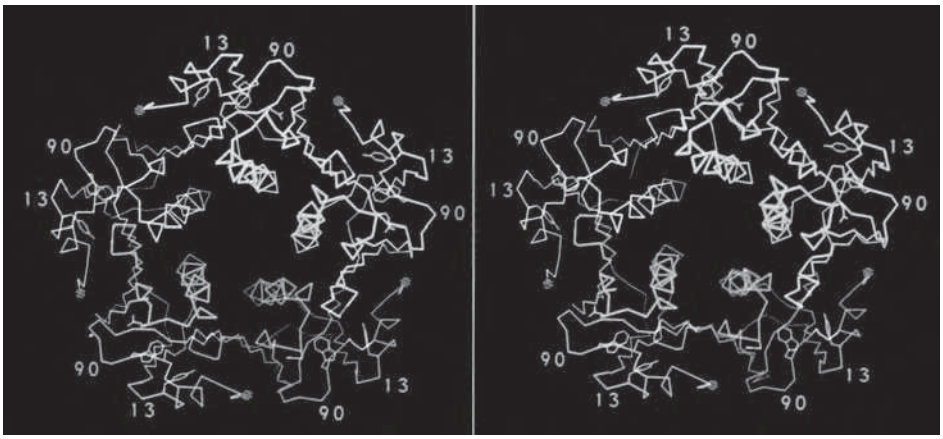



**Figure 7-5** Stereoscopic view of the $B_5$ pentamer of cholera toxin B. The pentamer, known as **choleragenoid**, has a central hole of ~1.5 nm diameter into which a helix from the A subunit is inserted. As viewed here, the front surface of the pentamer has binding sites for the oligosaccharide chains of ganglioside $GM_1$, which serves as the toxin receptor. The back side binds the A subunit. See also Box 11-A. From Zhang *et al.*[31]

Now consider what will happen to the little rings within the cell if the process that removes P to a non-associating form suddenly becomes more active so that [P] falls to $10^{-6}$ M. If $K$ is still $10^4$, what will be the percentages of P, $P_2$, and $P_3$ at equilibrium? Here we note a characteristic of cooperative processes: A higher than first power dependence on a concentration.

***Helical structures.*** If the angle at the interface *aj* is slightly different, instead of a closed ring, we obtain a helix as shown in Fig. 7-6A. The helix may have an integral number of subunits per turn or it may have a nonintegral number, as in the figure. The same type of heterologous interaction *aj* is involved in joining each subunit to the preceding one, but in addition other interactions occur. If the surfaces involved in these additional interactions are complementary and the geometry is correct, groups from two different parts of the molecule (e.g., *b* and *k*) may fit together to form another heterologous bond. Still a third heterologous interaction *cl* may be formed between two other parts of the subunit surfaces. If interactions *aj*, *bk*, and *cl* are strong (i.e., if the surfaces are highly complementary over large areas), extremely strong microtubular structures may be formed, such as those in the flagella of eukaryotic organisms (Fig. 1-8). If the interactions are weaker, labile microfilaments and microtubules, such as are often observed to form and dissociate within cells, may arise.

The geometry of subunits within a helix is often advantageously displayed by imagining that the surface of the structure can be unfolded to give a radial projection (Fig. 7-6B). Here subunits corresponding to those in the helix in Fig. 7-6A are laid out on a plane obtained by slitting the cylinder representing the surface of the helix and laying it out flat. In the example shown, the number of subunits per turn is about 4.8 but it can be an integral number. The interactions *bk* between subunits along the direction of the fiber axis may sometimes be stronger than those (*aj*) between adjacent subunits around the spiral. In such cases the microtubule becomes frayed at the ends through breaking of the *aj* interactions. This phenomenon can be observed under the electron microscope for the microtubules from flagella of eukaryotic organisms. Figures 7-7 to 7-10 show four helical structures from the molecular domain. They are a filamentous bacteriophage, a plant virus, a bacterial pilus, and an actin microfibril. Each is composed largely of a single kind of protomer. A larger and more complex helical structure, the microtubule, is shown in Fig. 7-34.

***Filamentous bacteriophages.*** Bacteriophages of the Ff family include the fd, f1, and M13 strains.[31a,32–36] Phage M13 is widely used in cloning genes and for many other purposes (Chapter 26). The genome is a circular, single-stranded DNA of ~6400 nucleotides

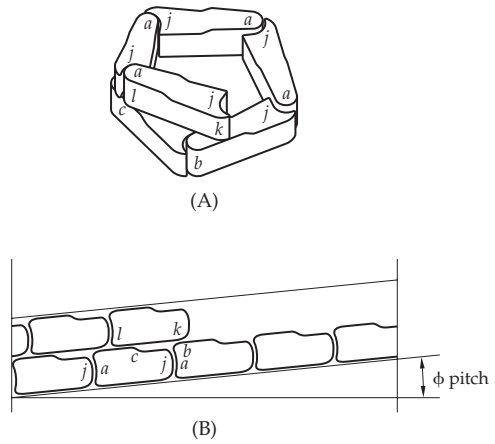



**Figure 7-6** (A) Heterologous bonding of subunits to form a helix. (B) Radial projection of subunits arranged as in helix A. Different bonding regions of the subunit are designated *a*, *b*, *c*, *j*, *k*, and *l*.

which is held in an elongated double-stranded form by a helical sheath of about 2700 subunits of a 50-residue protein. The rod is about 6 nm in diameter and 880 nm long and it is capped by two specialized proteins at one end and a different pair of proteins at the other end. The five coat proteins are encoded by five of the 11 genes present in these little viruses.[37,38]

Each coat subunit in the Ff viruses is coiled into an α-helical rod of 7 nm length. These are arranged in the virus in a right-handed helical pattern with a pitch of 1.5 nm and with 4.4 subunits per turn (Fig. 7-7). The protein rods are inclined to the helix axis and extend inward. This arrangement permits a "knobs-in-holes" hydrophobic bonding between subunits. The helix of pitch 1.5 nm is the **primary** or **one-start helix**. However, in every regular helical structure we can also trace a two-start helix, a three-start helix, etc. In this instance the five-start helix is easiest to see.

The protein coat of these viruses provides an elongated cylindrical cavity to protect the circular, single-stranded DNA molecule that is the genome. Although there are two antiparallel strands of DNA, a regular base-paired structure is impossible and the DNA is probably not present in a highly ordered form.[38a] There are about 2.4 nucleotides in the DNA per protein subunit. However, there are related viruses with ratios as low as one nucleotide per protein subunit and containing more highly extended DNA.[34,39]

***A rod-shaped plant virus.*** The tobacco mosaic virus (Figs. 5-41, 7-8) is a 300-nm-long rod constructed from 2140 identical wedge-shaped subunits whose detailed molecular structure is known.[40] Each 158-residue subunit contains five helices and a small β sheet. A single strand of RNA containing 6395 nucle-

otides (~ 3 per protein subunit) lies coiled in a groove where it interacts with side chains from two of the helices (Fig. 7-8B).[41–44a] The virus is assembled by the binding of a region of the RNA 800 – 1000 nucleotides from the 3' end to a two-turn helix of subunits that appears to form spontaneously. Additional subunits then add at each end, binding to the RNA as well as the adjacent protein subunits.[42,44,45] A relative with a very similar structure is cucumber green mottle mosaic virus.[46]

***Bacterial pili.*** The adhesion pili, or fimbriae,[48] of bacteria are also helical arrays of subunits. The P pili of *E. coli* are encoded by a cluster of 11 genes in the *pap* (pilus associated with pyelonephritis) cluster. They are needed to allow the bacteria to colonize the human urinary tract. The bulk of the ~1-μm-long pilus is made up of about 1000 subunits of a 185-residue protein encoded by gene *PapA*. They form a right-handed helix of ~7 nm diameter with 3.28 subunits per turn and a pitch of ~2.5 nm (Fig. 7-9A).[49–51] The rod is anchored to the bacterial outer membrane by a protein encoded by gene *PapH*, while subunits encoded by *PapE* and *PapF* fasten the adhesin protein (PapG) to the tips of the pili.[49,52,53,53a] The adhesin binds to the Galα1→4Gal ends of glycolipids in

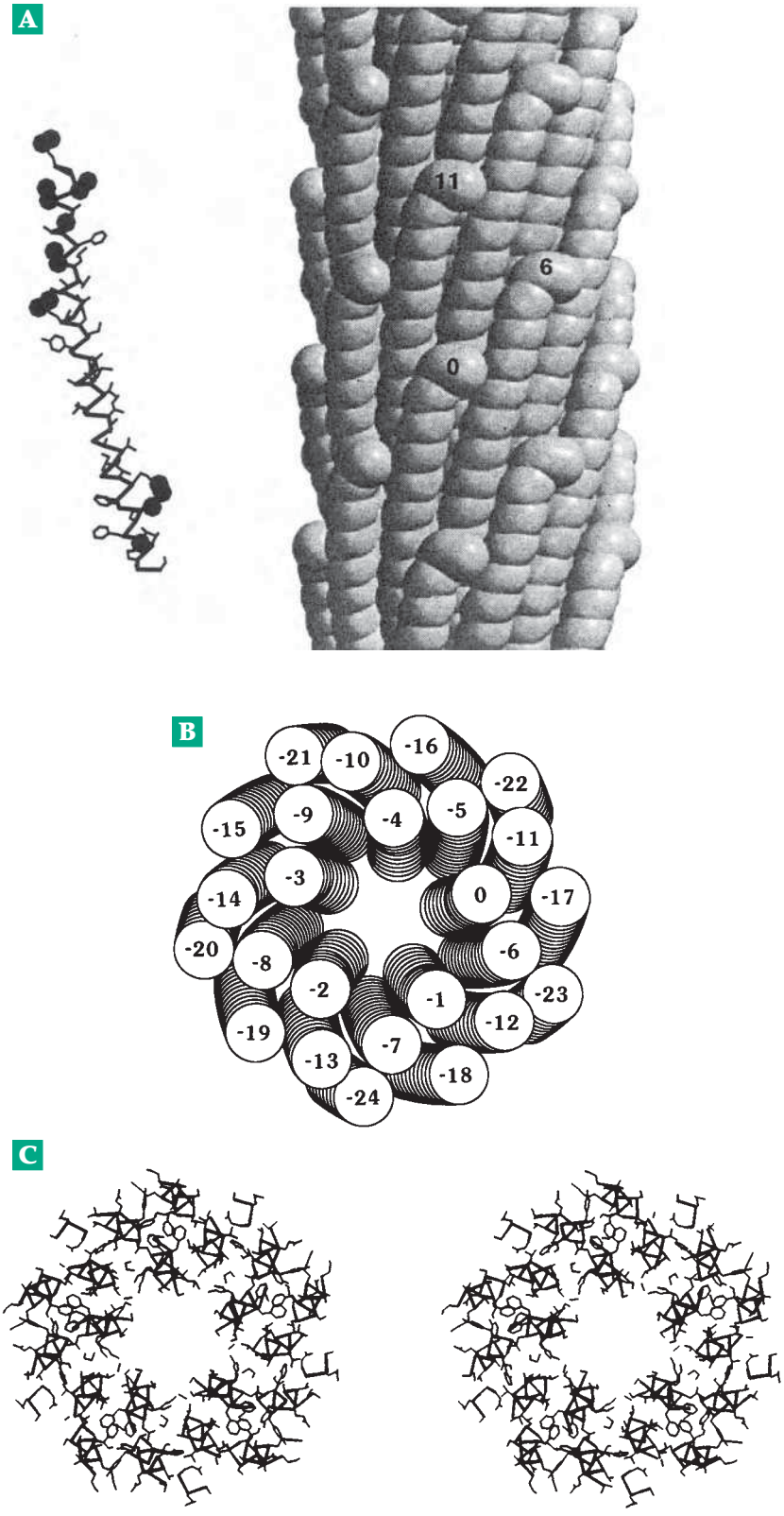



**Figure 7-7** Structure of the virus fd protein sheath. (A) Left. A single coat subunit, with its N terminus towards the top, as if moved from the left side of the sheath. The dark circles represent charged atoms of Asp, Glu, and Lys side chains. The backbone of the protein is a $C^{\alpha}$ diagram. The positively charged atoms near the C terminus line the inner surface of the sheath neutralizing the negative charge of the DNA core. Right. Each subunit is represented by a helical tube through successive $C^{\alpha}$ atoms. Three nearest neighbors, indexed as 0, 6, and 11, are indicated. The axial slab shown represents ~1% of the total length of the virion. From Marvin.[31a] (B) A 2.0 nm section through the virus coat with the helices shown as curved cylinders. The view is down the axis from the N-terminal ends of the rods. The rods extend upward and outward. The rods with indices 0 to –4 start at the same level, forming a five-start helical array. The rods with more negative indices start at lower levels and are therefore further out when they are cut in this section. (C) The same view but with "wire models" of the atomic structure of the rods. From Marvin *et al.*[32]

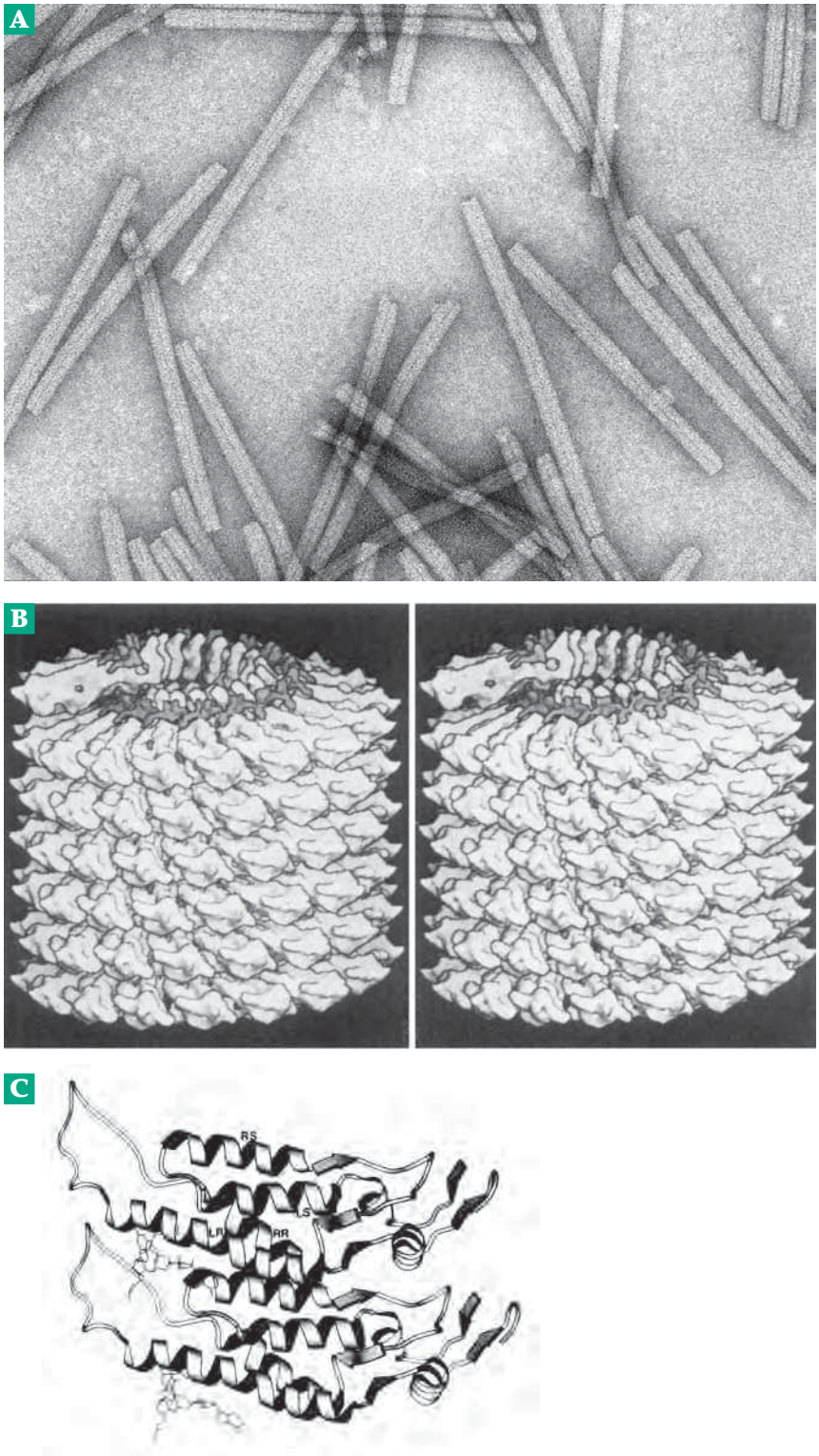



**Figure 7-8** (A) Electron micrograph of the rod-shaped particles of tobacco mosaic virus. © Omikron, Photo Researchers. See also Butler and Klug.[42] (B) A stereoscopic computer graphics image of a segment of the 300 nm long tobacco mosaic virus. The diameter of the rod is 18 nm, the pitch of the helix is 2.3 nm, and there are 16 1\3 subunits per turn. The coat is formed from ~2140 identical 17.5-kDa subunits. The 6395-nucleotide genomic RNA is represented by the dark chain exposed at the top of the segment. The resolution is 0.4 nm. From Namba, Caspar, and Stubbs.[47] (C) A MolScript ribbon drawing of two stacked subunits. From Wang and Stubbs.[46]

the kidney.[51,54] The PapE, F, and G subunits form a thin ~2.5-nm-thick by 15 nm "fibrillum" which is attached by an adapter protein encoded by gene *PapK*. A special chaperonin (*PapD* gene) is also required for pilus assembly in *E. coli*[53a] as well as other bacteria.[55] Another *E. coli* pilus adheres to mannose oligosaccharides.[53b]

Similar pili of *Neisseria gonnorheae* are used by that bacterium. The three-dimensional structure of the 158-residue pilin subunit is that of a globular subunit with an 8.5 nm α-helical spine at one end.[56,57] A proposed model of the intact pilus shaft is shown in Fig. 7-9B,C. Notice the similarity of the packing of the α-helical spines in the center to the packing arrangement in the bacteriophage coat in Fig. 7-7. Similar features may be present in the P pilus rod shown in Fig. 7-9A. However, there is uncertainty about the packing arrangement. The *E. coli* type 1 pilus subunits contain immunoglobulin folds that are completed by donation of an N-terminal strand from a neighboring subunit.[53a] In thin fimbriae of *Salmonella* extended, parallel β helices may be formed (see Fig. 2-17)[53e,53f]

Other types of pili are also well-known.[53c,d] F pili or conjugative pili are essential for sexual transfer of DNA between bacterial cells (Chapter 26). $F^+$ strains of *E. coli* form hollow pili of 8.5 nm diameter with a 2.0-nm central hole.[58,59] Their 90-residue subunits apparently form rotationally symmetric pentamers which stack to form the pili.[59] These pili are essential to establishing the initial contact between conjugating bacterial cells.

### *The thin filaments of muscle.*

An essential component of skeletal muscle (discussed further in Chapter 19) is filamentous **actin** (F-actin). It is composed of 375-residue globular subunits of a single type and with a highly conserved sequence.[60,61] It is found not only in muscle but also in other cells where it is a component of the cytoskeleton. The actin microfilament has the geometry of a left-

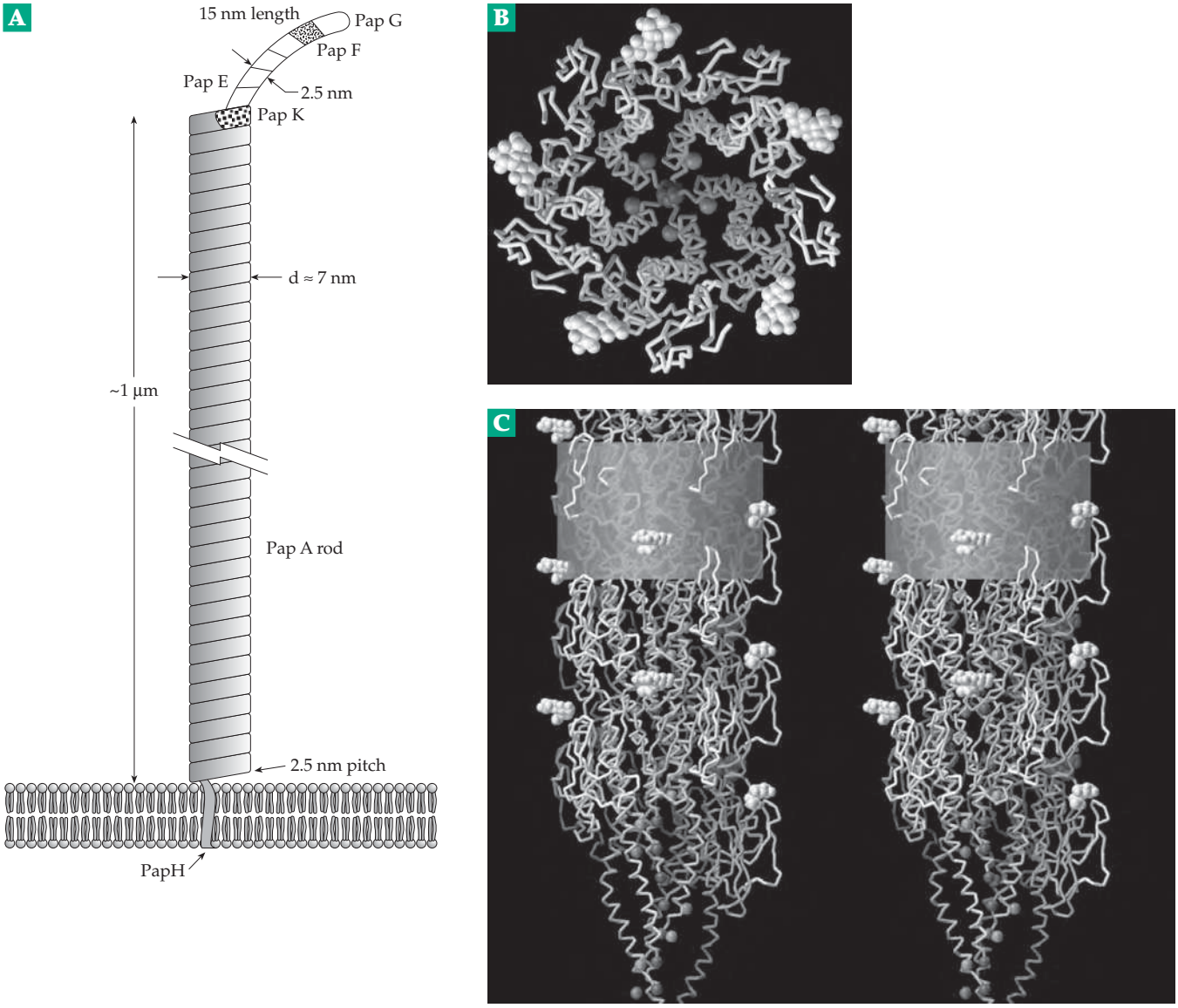



**Figure 7-9** (A) Schematic diagram of a bacterial P pilus. The ~1-µm-long helical rod is anchored to the outer cell membrane by protein Pap H. The adhesion Pap G binds to galactosyl glycolipids of the host. (B, C) The structure of pilus fiber from *Neisseria gonnorrheae* modeled from the atomic structure of the 54-residue pilin subunit. The exact structure of the fiber is uncertain, but the model generated here by trying various possible helical packings matches the dimensions obtained from fiber diffraction patterns and electron microscopic images. (B) Cross section. (C) Stereoscopic view. The experimental dimensions of 4.1-nm pitch and 6.0-nm diameters are shown by the "transparent" ring in (B). From Parge *et al.*[57]

handed one-start or **primary helix** with a pitch of only 0.54 nm and with approximately two subunits per turn (Fig. 7-10).[62,63]
It can also be described as a right-handed two-start helix in which *two chains* of subunits coil around one another with a long pitch (Fig. 7-10).

## 2. Oligomers with Twofold (Dyad) Axes

***Paired interactions.*** If two subunits are held together with interactions *aj* and are related by a twofold axis of rotation as shown in Fig. 7-11, we obtain an **isologous dimer**. Each point such as *a* in one subunit is related to the same point in the other subunit by reflection through the axis of rotation. In the center, along the twofold axis, points *c* and *c′* are *directly opposite the same points* in the other subunit. Figure 7-11

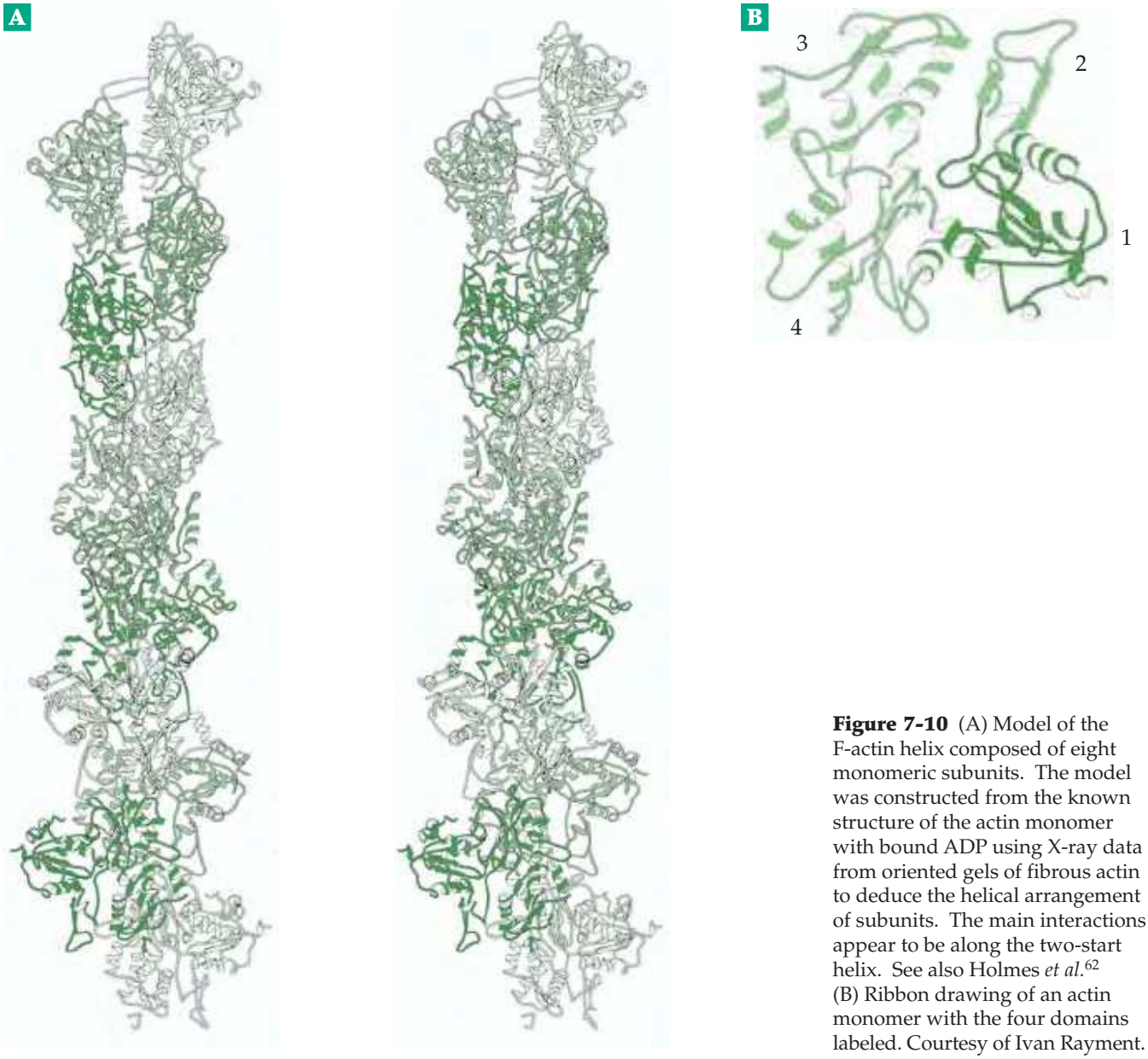



**Figure 7-10** (A) Model of the F-actin helix composed of eight monomeric subunits. The model was constructed from the known structure of the actin monomer with bound ADP using X-ray data from oriented gels of fibrous actin to deduce the helical arrangement of subunits. The main interactions appear to be along the two-start helix. See also Holmes *et al.*[62] (B) Ribbon drawing of an actin monomer with the four domains labeled. Courtesy of Ivan Rayment.

is drawn with a hole in the center so that groups *c* and *c′* do not actually touch, and it is the paired interactions such as *aj* of groups not adjacent to the axis that contribute most to the bonding. However, a real protein dimer may or may not have such a hole. The pair of identical interactions in an isologous dimer may be referred to as a single **isologous bond**. Such a bond always contains the paired interactions between complementary groups (*aj*) and has pairs of identical groups along the axis. However, because they are identical those groups usually cannot interact in a specific complementary manner.

Isologous bonding is very important in oligomeric enzymes, and it has been suggested that isologous interactions evolved early. Initially there may not have been much complementarity in the bonding but two "hydrophobic spots" on the surface of the subunits came together in a nonspecific association.[64] Later in evolution the more specific paired interactions could have been added.

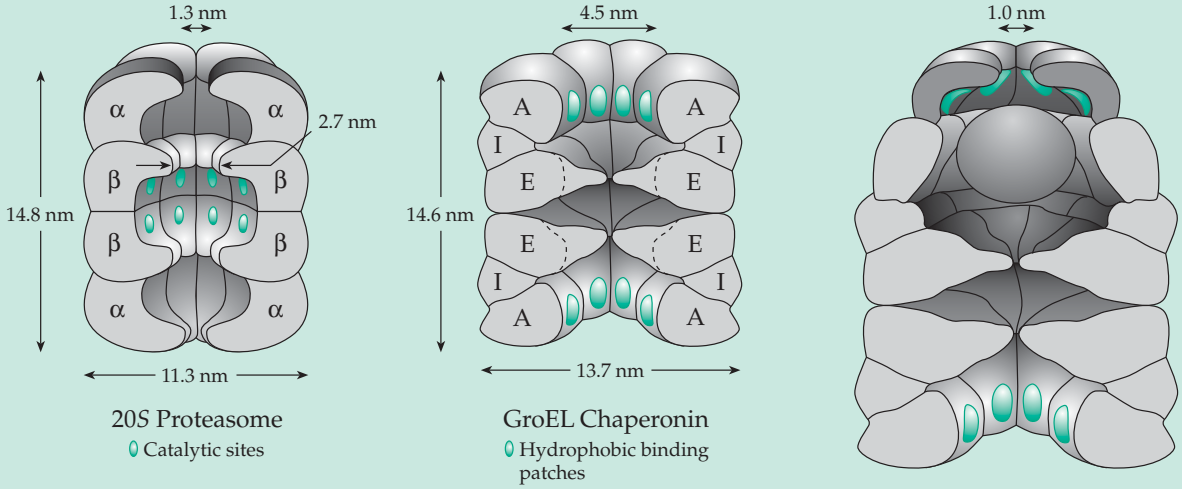
***Dihedral symmetry.*** Isologous dimers can serve as subunits in the formation of larger closed oligomers and helices; for example, an isologous pair of the sort shown in Fig. 7-11A can be flipped over onto the top of another similar pair as shown in Figs. 7-11B and 7-11C. Again, if the proper complementary surfaces exist, bonds can form as shown (*bk* in Fig. 7-11B and *bk* and *cl* in Fig. 7-11C). Both structures in Figs. 7-11B and 7-11C possess dihedral ($D_2$) symmetry.[65] In addition to the twofold axis of rotation lying perpendicular to the two rings, there are two other twofold axes of rotation as

## BOX 7-A LIFE AND DEATH FOR PROTEINS: CHAPERONINS AND PROTEASOMES

In 1968, a tiny cylindrical particle, which appeared to be a stack of 11-nm rings, was observed by electron microscopy of an extract of erythrocytes.[a,b] Later, a similar particle was found in both the nucleus and the cytoplasm of other cells of many organisms. The particles were soon recognized as a new type of protein-hydrolyzing enzyme, a large 700-kDa particle consisting of 20–30 subunits of several different types which came to be known as the **multicatalytic protease** or **20S proteasome**.[c,d] Electron microscopy and X-ray diffraction showed that the particle is formed from four stacked rings, each of which consists of seven subunits whose molecular masses range from 21 to 31 kDa.[e–j]

Proteasomes are strikingly similar in architecture, though not in peptide sequences, to another particle found in both bacteria and eukaryotes: a molecular "chaperone" or **chaperonin**. The chaperonins, of which there are several types, protect proteins while they fold or undergo translocation within cells.[k] One of the best studied members is the *E. coli* protein **GroEL**, which is also composed of double rings of 14 subunits with seven-fold rotational symmetry and with two of these assemblies associated back-to-back with dihedral symmetry.[l] The dimensions of GroEL and 20S proteasomes are nearly the same. However, GroEL has only two rings of ~60-kDa subunits, more than twice the size of proteosomal subunits. The accompanying sketch illustrates this fact and also the basic structural similarity of 20S proteosomes with GroEL. The αβ pairs of the proteosome, correspond to single subunits of the chaperonin, but these subunits have three distinct domains–apical, intermediate, and equatorial (labeled A, I, and E , respectively, in the drawing).[m] After a protein, whether correctly, incorrectly, or only partially folded, enters a cavity in GroEL, a second protein **GroES** of smaller size (~10 kDa) but with seven-fold symmetry binds to one end of the chaperonin.[p] Seven molecules of ATP also bind to sites on the GroEL ring to which GroES binds (the *cis* ring). The binding of the ATP and GroES evidently induces a major conformational change in the GroEL subunits[l,m,n,q] which causes the binding cavity to expand to over twice the original volume. This change (see drawing) also causes hydrophobic surfaces of the cavity to become buried and hydrophilic side chains to be exposed. The cavity surface was initially largely hydrophobic and able to bind many proteins nonspecifically, but upon expansion it becomes hydrophilic and less likely to bind. This releases the encased protein to complete its folding or to partially unfold and refold without interference from other proteins.

While a protein is adjusting its folding in the *cis* compartment another protein molecule may become trapped in the *trans* compartment. After some time the bound ATP molecules are hydrolyzed. As in the contraction of muscle, which is discussed in Chapter 19, the loss of inorganic phosphate ($P_i$) and ADP from the active site can be accompanied by movement. In the chapenonin this involves a conformational switch so that the ES heptamer is released and the conformation of the *trans* ring of EL is switched to that of the initial *cis* ring and vice versa. The new *cis* ring is ready to receive an ES cap and the new *trans* ring can release the folded protein.[r] A variety of experimental approaches are being used in an




From Weissman, Sigler and Horwich[g]. Illustrations by K. Sutliff.
Right: From Mayhew and Hartl[o]

Expanded GroEL-GroES complex with enclosed folding polypeptide[o]

## BOX 7-A LIFE AND DEATH FOR PROTEINS: CHAPERONINS AND PROTEASOMES (continued)

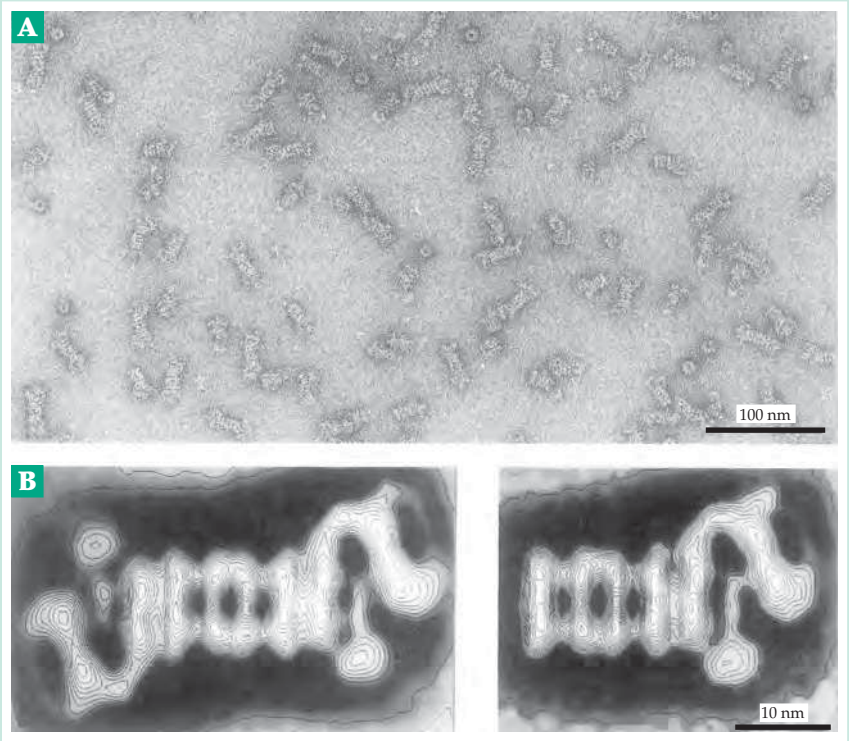
effort to further understand the action of GroEL.[s–x] The chaperonin may function repeatedly before a protein becomes properly folded.[t]

While chaperonins assist proteins to fold correctly proteasomes destroy unfolded chains by partial hydrolysis, cutting the chains into a random assortment of pieces from 3 to 30 residues in length with an average length of ~8 residues.[y] Proteasomes destroy not only unfolded and improperly folded proteins but also proteins marked for destruction by the ubiquitin system described in Box 10-C. It has been hard to locate true proteosomes in most bacteria. However, they do contain protease particles with similar characteristics[z–bb] and archaeons, such as *Thermoplasma acidophilum*, have proteasomes similar to those of eukaryotes.[cc]

The *Thermoplasma* proteasome contains only two kinds of subunits, α and β, which have similar amino acid sequences. These form $\alpha_7$ and $\beta_7$ rings which associate in $\alpha_7\beta_7$ pairs with two of these double rings stacked back-to-back with dihedral D7 symmetry: $\alpha_7\beta_7\beta_7\alpha_7$. The crystal structure has been determined for this 20S proteasome from *T. acidophilum*[g,dd] and for the corresponding proteasome from yeast (*Saccharomyces cerevisiae*).[f,ee] The accompanying drawings illustrate top and side views of the *T. acidophilum* proteasome. The particle contains three internal cavities. The outer two are formed between the $\alpha_7$ and $\beta_7$ rings and the inner is formed between the two $\beta_7$ rings. A channel only 1.3 nm in diameter permits the entrance of peptide chains into the compartments.

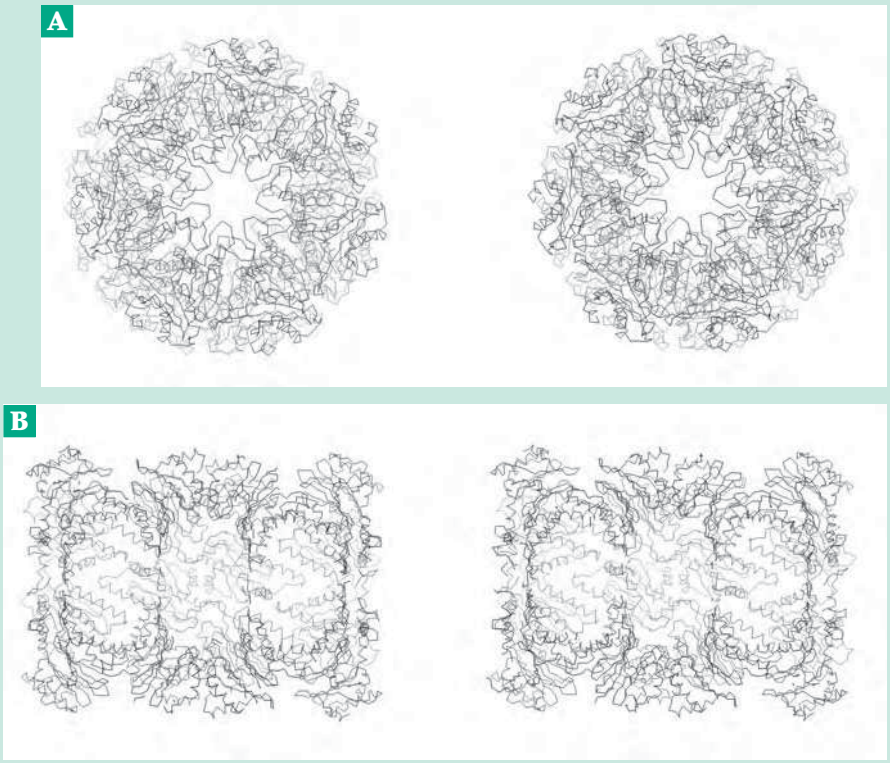
The active sites of the enzymes[ff] are located in the β subunits in the central cavity.[dd] While the yeast and human proteasomes are similar to those of *Thermoplasma*, the β subunits consist of seven different protein-hydrolyzing enzymes whose catalytic activities and mechanisms are considered in Chapter 12. There are also seven different α subunits, all of whose sequences are known.[i,gg] To make the story more complex, additional subunits, some of which catalyze ATP hydrolysis, form a 600- to 700-kDa cap which adds to one or both ends of a 20S proteosome to give a larger **26S proteasome**.[b,d,e,w] These larger proteasomes carry out an ATP-dependent cleavage of proteins selected for degradation by the ubiquitin system (Box 10-C; Chapter 12). Some of the short peptide segments formed by proteasomes may leave cells and participate in intercellular communication. For example, pieces of antigenic peptides are used by cells of the immune system for "antigen presentation" (Chapter 31),[hh] an important process by which the immune system recognizes which cells are "self" and which are foreign or malignant and must be killed.

The structure of the caps on the 26S proteasome ends is complex. At least 20 different regulatory subunits have been identified.[ii,jj]




(A) Electron micrograph of 26S proteasomes from *Xenopus* oocytes negatively stained with 2% uranyl acetate. (B) Image of the 26S proteasome (left) and a 20S proteasome with only one end cap. These views were obtained by correlation averaging of 527 individual images of the 26S proteasome and 395 images of the single-ended form. From Peters *et al.*[e] Courtesy of Wolfgang Baumeister.

## BOX 7-A (continued)




(A) Top view of the 20S proteasome as an α-carbon plot showing the seven-fold symmetry. The α subunits are in front of the β subunits. (B) Side view showing the proteasome cut open along its seven-fold axis. From Löwe *et al.*[dd] Courtesy of Robert Huber.


[a] Harris, J. R. (1968) *Biochim. Biophys. Acta.* **150**, 534–537
[b] Peters, J.-M. (1994) *Trends Biochem. Sci.* **19**, 377–382
[c] Bosch, G., Baumeister, W., and Essen, L.-O. (2000) *J. Mol. Biol.* **301**, 19–25
[d] Ferrell, K., Wilkinson, C. R. M., Dubiel, W., and Gordon, C. (2000) *Trends Biochem. Sci.* **25**, 83–88
[e] Peters, J.-M., Cejka, Z., Harris, J. R., Kleinschmidt, J. A., and Baumeister, W. (1993) *J. Mol. Biol.* **234**, 932–937
[f] Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H. D., and Huber, R. (1997) *Nature (London)* **386**, 463–471
[g] Weissman, J. S., Sigler, P. B., and Horwich, A. L. (1995) *Science* **268**, 523–524
[h] DeMartino, G. N., and Slaughter, C. A. (1999) *J. Biol. Chem.* **274**, 22123–22126
[i] Gerards, W. L. H., de Jong, W. W., Bloemendal, H., and Boelens, W. (1998) *J. Mol. Biol.* **275**, 113–121
[j] Schmidtke, G., Schmidt, M., and Kloetzel, P.-M. (1997) *J. Mol. Biol.* **268**, 95–106
[k] Netzer, W. J., and Hartl, F. U. (1998) *Trends Biochem. Sci.* **23**, 68–73
[l] Sigler, P. B., Xu, Z., Rye, H. S., Burston, S. G., Fenton, W. A., and Horwich, A. L. (1998) *Ann. Rev. Biochem.* **67**, 581–608
[m] Kawata, Y., Kawagoe, M., Hongo, K., Miyazaki, T., Higurashi, T., Mizobata, T., and Nagai, J. (1999) *Biochemistry* **38**, 15731–15740
[n] Betancourt, M. R., and Thirumalai, D. (1999) *J. Mol. Biol.* **287**, 627–644
[o] Mayhew, M., and Hartl, F. U. (1996) *Science* **271**, 161–162
[p] Hunt, J. F., Weaver, A. J., Landry, S. J., Gierasch, L., and Deisenhofer, J. (1996) *Nature (London)* **379**, 37–45
[q] Cliff, M. J., Kad, N. M., Hay, N., Lund, P. A., Webb, M. R., Burston, S. G., and Clarke, A. R. (1999) *J. Mol. Biol.* **293**, 667–684
[r] Rye, H. S., Burston, S. G., Fenton, W. A., Beechem, J. M., Xu, Z., Sigler, P. B., and Horwich, A. L. (1997) *Nature (London)* **388**, 792–798
[s] Chatellier, J., Buckle, A. M., and Fersht, A. R. (1999) *J. Mol. Biol.* **292**, 163–172
[t] Nieba-Axmann, S. E., Ottiger, M., Wüthrich, K., and Plückthun, A. (1997) *J. Mol. Biol.* **271**, 803–818
[u] Gervasoni, P., Gehrig, P., and Plückthun, A. (1998) *J. Mol. Biol.* **275**, 663–675
[v] Torella, C., Mattingly, J. R., Jr., Artigues, A., Iriarte, A., and Martinez-Carrion, M. (1998) *J. Biol. Chem.* **273**, 3915–3925
[w] Horwich, A. L., Weber-Ban, E. U., and Finley, D. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11033–11040
[x] Buckle, A. M., Zahn, R., and Fersht, A. R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3571–3575
[y] Kisselev, A. F., Akopian, T. N., and Goldberg, A. L. (1998) *J. Biol. Chem.* **273**, 1982–1989
[z] Kessel, M., Maurizi, M. R., Kim, B., Kocsis, E., Trus, B. L., Singh, S. K., and Steven, A. C. (1995) *J. Mol. Biol.* **250**, 587–594
[aa] Shin, D. H., Lee, C. S., Chung, C. H., and Suh, S. W. (1996) *J. Mol. Biol.* **262**, 71–76
[bb] Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F., and Baumeister, W. (1996) *Science* **274**, 1385–1389
[cc] Maupin-Furlow, J. A., and Ferry, J. G. (1995) *J. Biol. Chem.* **270**, 28617–28622
[dd] Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995) *Science* **268**, 533–539
[ee] Stuart, D. I., and Jones, E. Y. (1997) *Nature (London)* **386**, 437–438
[ff] Voges, D., Zwickl, P., and Baumeister, W. (1999) *Ann. Rev. Biochem.* **68**, 1015–1068
[gg] Schmidt, M., and Kloetzel, P.-M. (1997) *FASEB J.* **11**, 1235–1243
[hh] Goldberg, A. L., and Rock, K. L. (1992) *Nature (London)* **357**, 375–379
[ii] Adams, G. M., Falke, S., Goldberg, A. L., Slaughter, C. A., DeMartino, G. N., and Gogol, E. P. (1997) *J. Mol. Biol.* **273**, 646–657
[jj] Knowlton, J. R., Johnston, S. C., Whitby, F. G., Realini, C., Zhang, Z., Rechsteiner, M., and Hill, C. P. (1997) *Nature (London)* **390**, 639–643

indicated in the drawings. Again, the interactions are paired and isologous; of many possible contacts two *bk* interactions and two *cl* interactions are marked for each pair of subunits in Fig. 7-11. There are a total of six pairs of these interactions, one between each combination of two subunits. This may be a little more difficult to see in Fig. 7-11B than in Fig. 7-11C because in the former the subunits are arranged in a more or less square configuration. Nevertheless, a pair of interactions between the left-hand subunit in the top ring and the subunit in the lower ring at the right does exist, even if it is only electrostatic and at a distance. An example of a tetrameric enzyme with perfect dihedral symmetry of the type shown in Fig. 7-11B is **lactate dehydrogenase** (Chapter 15). The plant agglutinin **concanavalin A** has a quaternary structure resembling that in Fig. 7-11C.

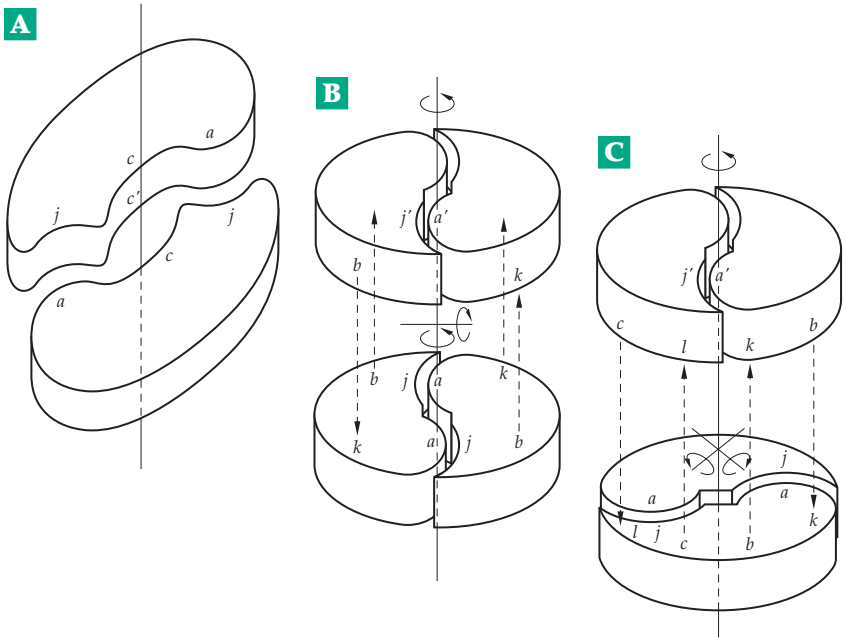



**Figure 7-11** (A) Isologous bonding between pairs of subunits; (B) an "isologous square" arrangement of subunits; (C) an apparently "tetrahedral" arrangement of subunits. Note the three twofold axes.

Square arrays of four subunits can be formed using either heterologous or isologous interactions. Both types of bonding can occur in larger aggregates. For example, two trimers such as that shown in Eq. 7-25 can associate to a hexamer having dihedral ($D_3$) symmetry; a heterologous "square tetramer" can dimerize to give a dihedral ($D_4$) octamer.[65] The enzymes **ornithine decarboxylase** (Fig. 7-12)[66] and **glutamine synthetase** (Chapter 24)[67] each consist of double rings of six subunits each. The upper ring is flipped over onto the lower giving dihedral symmetry ($D_6$) with one 6-fold axis and six 2-fold axes at right angles to it.

**Figure 7-12** A ribbon representation of the ornithine decarboxylase dodecamer. Six dimers of the 730-residue subunits are related by $C_6$ crystallographic symmetry. MolScript drawing from Momany *et al.*[66] Courtesy of Marvin Hackert.

### *Oligomers with cubic symmetry (polyhedra).*

Symmetrical arrangements containing more than one axis of rotation of order higher than 2-fold are said to have cubic symmetry. The **tetrahedron** is the simplest example. It contains four 3-fold axes which pass through the vertices and the centers of the faces and three 2-fold axes which pass through the midpoints of the six edges. *Since protein subunits are always asymmetric, a tetrameric protein cannot possess cubic symmetry*. As we have already seen, tetrameric enzymes have dihedral symmetry. However, a heterologous trimer with 3-fold symmetry can form a face of a tetrahedron containing a total of 12 asymmetric subunits. Twenty-four subunits can interact to form a **cube**. Three 4-fold axes pass through the centers of the faces, four 3-fold axes pass through the vertices, and six 2-fold axes pass through the edges (see Figs. 7-13 and 16-3).

The largest structure of cubic symmetry that can

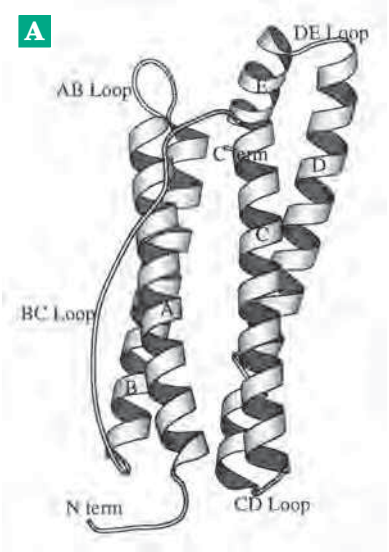




**Figure 7-13** (A) MolScript ribbon drawing of a subunit of the iron oxide storage proteins L-ferritin from amphibian red cells. This 4-helix bundle is represented by cylinders of 1.3 nm diameter in the oligomer. (B) Helices A and C of the monomer are on the outer surfaces of the oligomer and helices B and D are on the inner surface. The oligomer consists of a shell of 24 subunits and is viewed down a 4-fold axis illustrating its 423 (cubic) symmetry. The molecule is illustrated further in Fig. 16-3. From Trikha *et al.*[74] Courtesy of Elizabeth Theil.

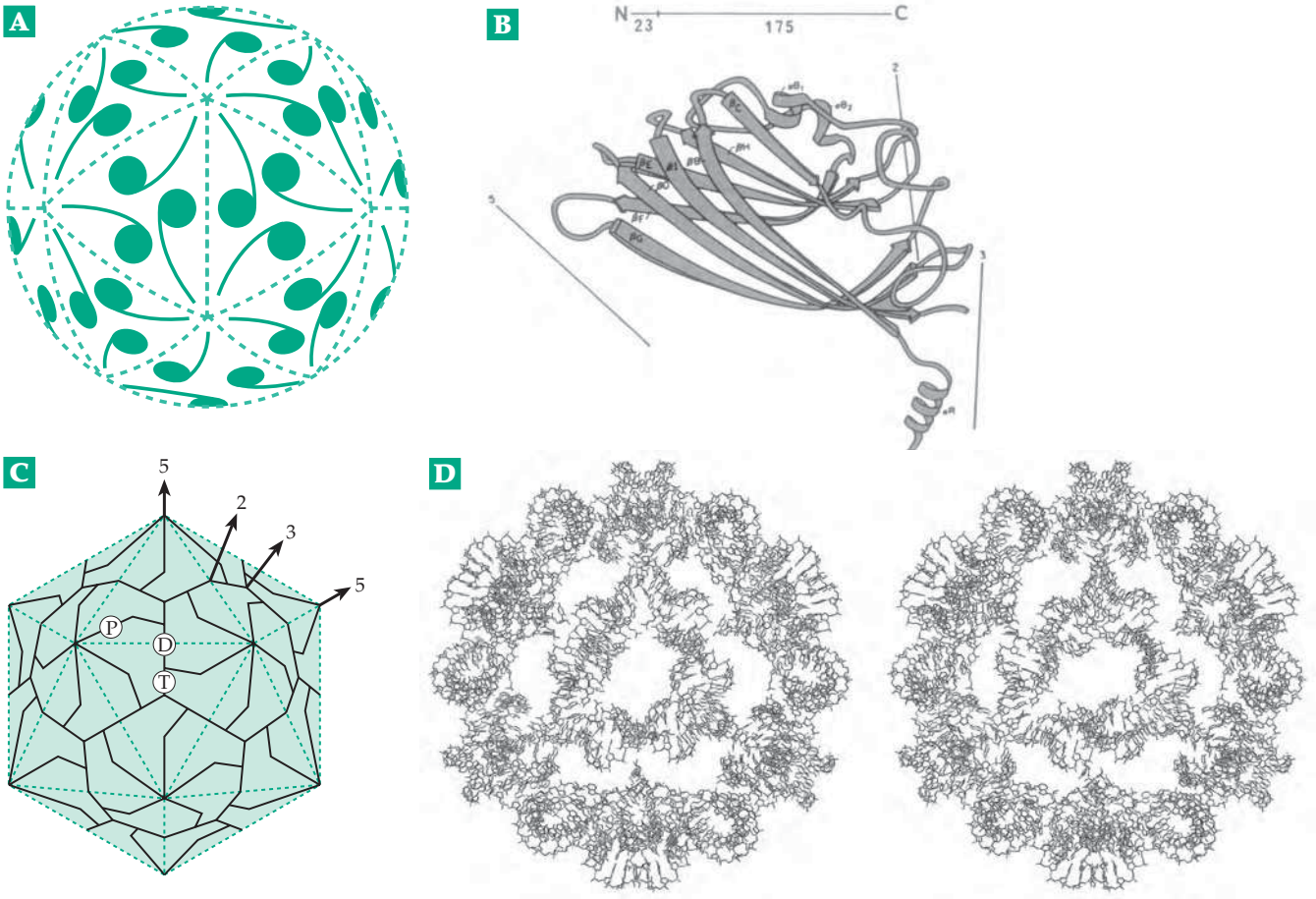



**Figure 7-14** (A) Schematic drawing illustrating an icosahedrally symmetric structure with sixty identical asymmetric subunits all in equivalent positions. The 5-fold axes are located at the vertices of the icosohedron and the 2-fold and 3-fold axes can readily be seen. (B) Ribbon drawing of the 195-residue polypeptide chain of the coat subunit of satellite tobacco necrosis virus. The protein folds into an inwardly projecting N-terminal segment and a "β-jellyroll" domain. The packing of this subunit in the virus particle is shown schematically in (C). The symmetry axes drawn next to the subunit diagram (B) can be used to position it in the structure. Contacts between subunits are labeled D, T, and P ('dimer', 'trimer', 'pentamer'). Diagrams courtesy of Drs. Strandberg, Liljas, and Harrison.[68] (D) The distribution of RNA helical segments in a hemisphere of a virion of a similar small virus, the satellite tobacco mosaic virus. The virion viewed down a 3-fold axis from the virus exterior. The helical axes of the RNA segments are along icosahedral edges. From Larson *et al.*[75]

be made is the **icosahedron**, a regular solid with 20 triangular faces. Sixty subunits, or some multiple of 60, are required and at each vertex they form a heterologous pentamer. As with the tetrahedron, each face contains a heterologous trimer, while isologous bonds across the edges form dimers (Fig. 7-14C). Many viruses consist of roughly spherical protein shells (coats) containing DNA or RNA inside.[68–70] As with the filamentous viruses, the protein coats consist of many identical subunits, a fact that can be rationalized in terms of economy from the genetic viewpoint. *Only one gene is needed to specify the structure of a large number of subunits.*[70,71] Under the electron microscope the viruses often have an icosahedral appearance (Figs. 5-41A, 7-14), and chemical studies show that the number of the most abundant subunits is usually a multiple of 60. An example is the tiny **satellite tobacco necrosis virus**,[72] diameter ~18 nm, whose coat contains just 60 subunits of a 195-residue protein. Its genome is a 1239-nucleotide molecule of RNA. The structure of the coat has been determined to 0.25-nm resolution.[73]

Many virus coats have 180 subunits or a number that is some other multiple of 60. However, in these coats the subunits cannot all be in identical environments. Two cases may be distinguished. If all of the subunits have identical amino acid sequences they probably exist in more than one distinct conformation that permit them to pack efficiently. (Next section) Alternatively, two or more subunits of differing sequence and structure may associate to form 60 larger subunits that do pack with icosahedral symmetry. For example, the polioviruses (diameter 25 nm) contain three major coat proteins (α, β, and γ or VP1, VP2, and VP3). These are formed by cleavage of a large precursor protein into at least four pieces.[76,77] The three largest pieces of ~33-, 30-, and 25-kDa mass (306, 272, and 238 residues, respectively) aggregate as $(\alpha\beta\gamma)_{60}$. Sixty copies of a fourth subunit of 60 residues are found within the shell. Related picorna viruses such as human rhinoviruses (Fig. 7-15),[69,78,79] foot-and-mouth disease virus, parvovirus,[80] and Mengo virus[81] have similar architectures. The small (diameter 25 nm) single-stranded DNA bacteriophages such as ϕX174 also have three different coat proteins, one of which forms small hollow spikes at the vertices of the icosahedral shell (Fig. 5-41A).[82]

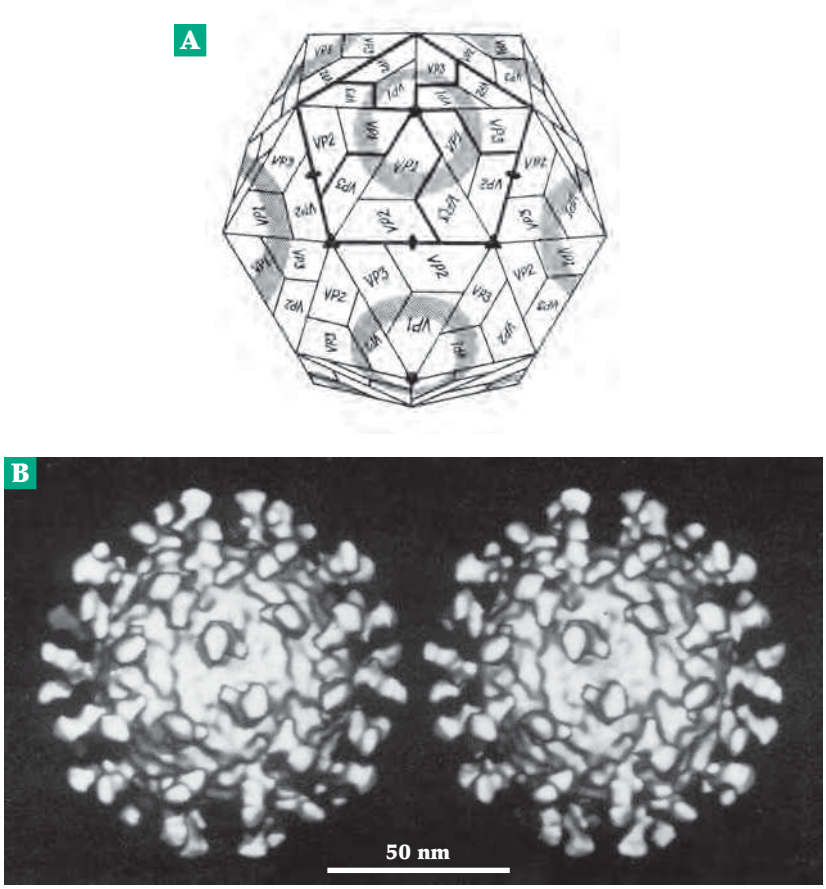



**Figure 7-15** (A) Schematic diagram of the icosahedral shell of a human rhinovirus showing the arrangement of the three subunits VP1, VP2, and VP3, each present as 60 copies. (B) Stereoscopic view of an image of the virus "decorated" by the binding of two immunoglobulinlike domains of the intercellular adhesion molecule ICAM-1, a natural receptor for the virus. Part of this receptor binds into a groove or "canyon," which in marked in (A) by the dark bands. From Olson *et al.*[78] Courtesy of Michael Rossmann.

***Asymmetry and quasi-equivalence in oligomers.*** It is natural to think about association of subunits in symmetric ways. Consequently the observation of square, pentagonal, and hexagonal arrangements of subunits directly with the electron microscope led to a ready acceptance of the idea that protomers tend to associate symmetrically. However, consider the predicament of the two molecules shown in Fig. 7-16. They might get together to form an isologous dimer if it were not for the fact that their "noses" are in the way. Despite the obvious steric hindrance, an isologous dimer can be formed in this case if one subunit is able to undergo a small change in conformation (Fig. 7-16). In the resulting dimer the two subunits are only **quasi-equivalent**.

Unsymmetrical dimerization of proteins appears to be a common phenomenon that is often observed in protein crystals. For example, the

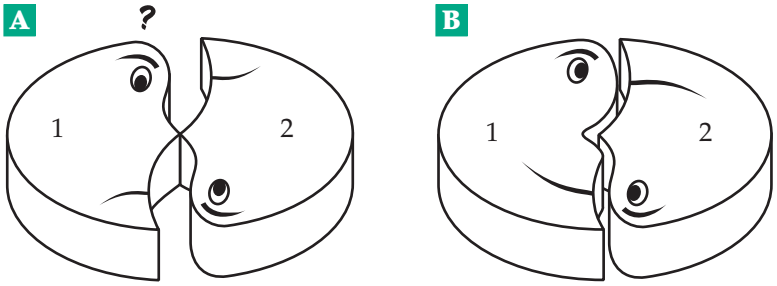



**Figure 7-16** Nonsymmetric bonding in a dimer. (A) Two molecules which cannot dimerize because of a bad fit at the center. (B) A solution: Molecule 1 has refolded its peptide chain a little, changing shape enough to fit to molecule 2.

enzymes malic dehydrogenase and glyceraldehyde phosphate dehydrogenase (Chapter 15) are both tetramers of approximate dihedral symmetry but X-ray crystallography revealed distinct asymmetries[83,84] which include a weaker binding of the coenzyme $NAD^+$ in one subunit. This may simply reflect differences in environment within the crystal lattice. However, negative cooperativity in coenzyme binds has also been revealed by kinetic experiments.[85]

The polypeptide hormone **insulin** is a small protein made up of two chains (designated A and B) which are held together by disulfide bridges (Fig. 7-17A). Figure 7-17B is a sketch of the structure as revealed by X-ray crystallography,[86,87] with only the backbone of the peptide chains and a few side chains shown. In the drawing, the B chain lies behind the A chain. Beginning with the N-terminal Phe 1 of the B chain the peptide backbone makes a broad curve, and then falls into an helix of three turns lying more or less in the center of the molecule. After a sharp turn, it continues upward on the left side of the drawing in a nearly completely extended β structure. The A chain has an overall U shape with two roughly helical portions. The U shape is partly maintained by a disulfide bridge running between two parts of the A chain. Two disulfide bridges hold the A and B chains together, and hydrophobic bonding of internal side chain groups helps to stabilize the molecule.

Insulin in solution dimerizes readily, the subunits occupying quasi-equivalent positions. Figure 7-17C shows some details of the bonding between the subunits in the insulin dimer with a view from the outside of the molecule down the 2-fold axis (marked by the X in the center of the Phe 25 ring in the right-hand chain) through the dimer. The C-terminal ends of the B chains are seen in an extended conformation. The two antiparallel chains form a β structure with two pairs of hydrogen bonds. If there were perfect isologous bonding, the two pairs would be entirely equivalent and symmetrically related one to the other. A straight line drawn from a position in one chain and passing through the twofold axis (X) would also pass approximately through the corresponding position in the other chain. However, there are many deviations from perfect symmetry, the most striking of which is at the center where the Phe 25 from the right-hand chain projects upward and to the left. If the symmetry were perfect the corresponding side chain from the left-hand chain would project upward and to the right and the two phenylalanines would collide, exactly as do the "noses" in Fig. 7-16. In insulin one phenylalanine side chain has been flipped back out of the way.

Under proper conditions, three insulin dimers associate to form a hexamer of approximate dihedral ($D_3$) symmetry that is stabilized by the presence of two zinc ions. Figure 7-17D is a crude sketch of the hexamer showing the three dimers, the 3-fold axis of symmetry, and the two pseudo 2-fold axes, one passing between the two subunits of the dimer and the other between two adjacent dimers. Figure 7-18 is a stereoscopic ribbon diagram of the atomic structure, with the A chains omitted, as obtained by X-ray diffraction.[87] The structure has also been obtained by NMR spectroscopy.[88] Note that each of the two zinc atoms lies on the threefold axis and is bound by three imidazole rings from histidines B-10. The significance of the zinc binding is uncertain but these hexamers readily form rhombohedral crystals, even within the pancreatic cells that synthesize insulin. The structure illustrates a feature that is common to many oligomers of circular or dihedral symmetry. A central "channel" is often quite open and protruding side chain groups, such as the imidazole groups in insulin, form handy nests into which ions or molecules regulating activity of proteins can fit. Conformational differences in insulin are induced by the binding of phenol. In Fig. 7-18A the C-terminal ends of the chains are extended but in the phenol complex (B) they have coiled to extend the α helices.

***Quasi-equivalence in virus coats.*** A large number of icosahedral viruses have coats consisting of 180 identical subunits. For example, the small RNA-containing bacteriophage MS2 consists of an eicosahedral shell of 180 copies of a 129-residue protein that encloses one molecule of a 3569-residue RNA.[89] The virus also contains a single molecule of a 44-kDa protein, the A protein, which binds the virus to a bacterial pilus to initiate infection. Related bacteriophages GA, fr, f2, and Qβ[90,91] have a similar

architecture. Many RNA-containing viruses of plants also have 180 subunits in their coats.[68] Much studied are the **tomato bushy stunt virus** (diameter ~33 nm, 40-kDa subunits),[68] and the related **southern bean mosaic virus**.[92] The human **wart virus** (diameter ~ 56 nm) contains 420 subunits, seven times the number in a regular icosahedron. **Adenoviruses** (diameter ~100 nm) have 1500 subunits, 25 times more than the 60 in a regular icosahedron.[93,94] Caspar and Klug[95] proposed a theory of quasi-equivalence of subunits according to which the distances between the centers of subunits are preserved in a family of **icosadeltahedra** containing subunits in multiples of 20. However, the angles must vary somewhat from those in a regular icosahedron (compare with geodesic shells in which the angles are constant but the distances are not all the same). The resulting polyhedra contain hexamers as well as pentamers at vertices; for example, the shells of the 180-subunit viruses contain clusters of subunits forming 12 pentamers and 20 hexamers. There are

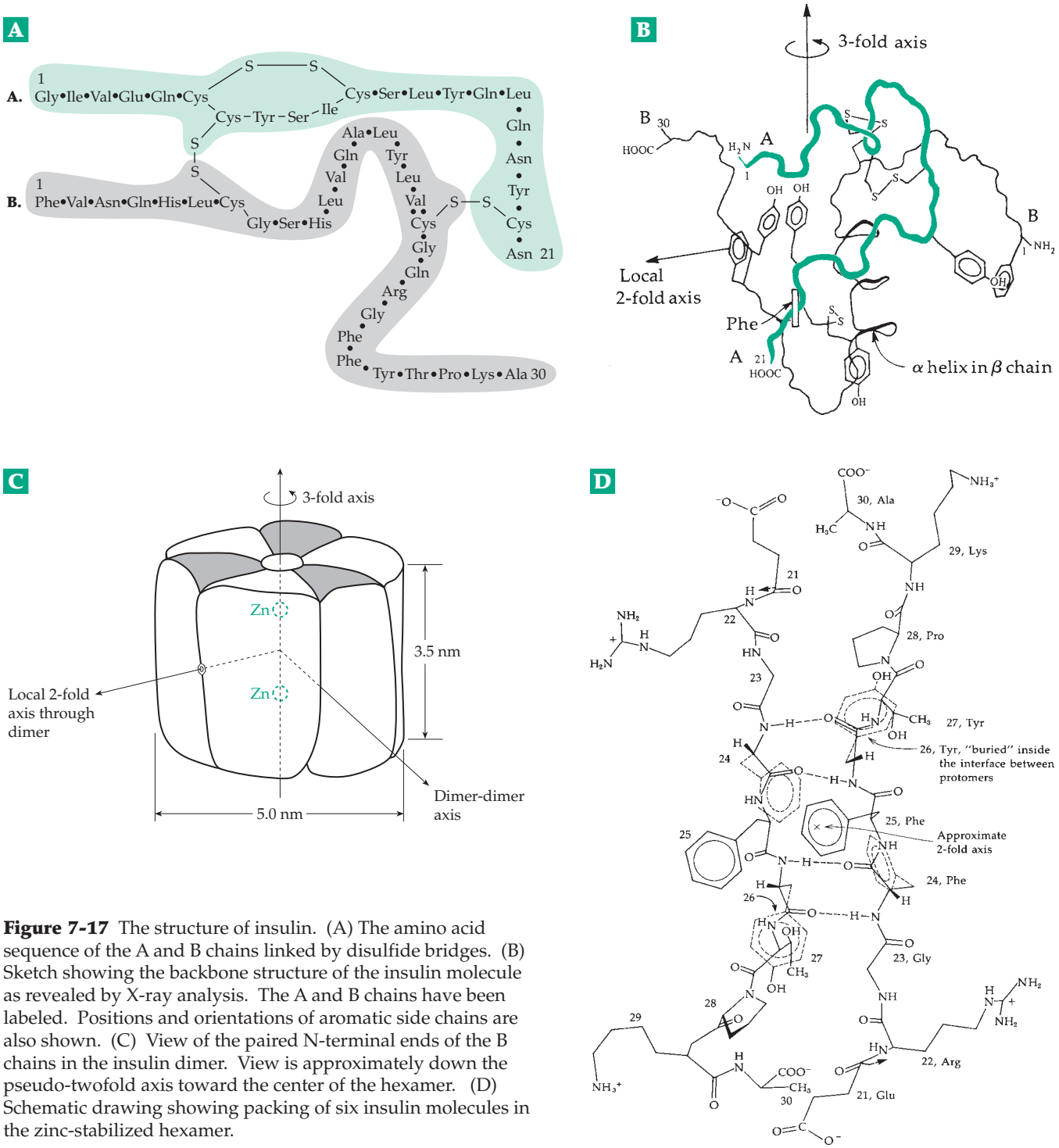



**Figure 7-17** The structure of insulin. (A) The amino acid sequence of the A and B chains linked by disulfide bridges. (B) Sketch showing the backbone structure of the insulin molecule as revealed by X-ray analysis. The A and B chains have been labeled. Positions and orientations of aromatic side chains are also shown. (C) View of the paired N-terminal ends of the B chains in the insulin dimer. View is approximately down the pseudo-twofold axis toward the center of the hexamer. (D) Schematic drawing showing packing of six insulin molecules in the zinc-stabilized hexamer.

also 60 trimers (on the faces) and 90 dimers (across the edges) (Fig. 7-19). Such structures can be formed only for certain values of $T$ where the number of subunits is $60T$ and there are 12 pentamers (pentons) and 10 ($T$-1) hexamers hexons. $T$ can assume values of $h^2 + hk + k^2$ where h and k may be positive integers or zero. Some allowed $T$ values are 1, 3, 4, 7, 9, 13, 25.[68,70,96,97]

The subunits in virus coats with $T$ greater than one are not all in equivalent positions. For example, the three subunits labeled A, B, and C in Fig. 7-19 are each slightly differently positioned with respect to neighboring subunits. Since virus coats are usually tightly packed the subunits must assume more than one conformation. One kind of conformational change that allows quasi-equivalence of subunits is observed in the tomato bushy stunt virus. Two structural domains are connected by a hinge which allows an outer protruding domain to move slightly to preserve good isologous interactions with a corresponding domain in another subunit.[68]

The southern bean mosaic virus has an eight-stranded antiparallel β-barrel structure closely similar to that of the major domain of the bushy stunt viruses but lacking the second hinged domain. The problem of quasi-equivalence is resolved by the presence of an N-terminal extension that binds onto a subunit across the quasi-six-fold axis to give a set of three subunits (labeled C in Fig. 7-19) that associate with true three-fold symmetry and another set (B) with a slightly different conformation fitting between them.[68,92] The subunits A, which have a third conformation, fit together around the five-fold axis in true cyclic symmetry.

A surprising finding is that the polyoma virus coat, which was expected to contain 420 (7 x 60) subunits, apparently contains only 360. The result is that the hexavalent morphological unit is a pentamer and that quasi-eqivalence appears

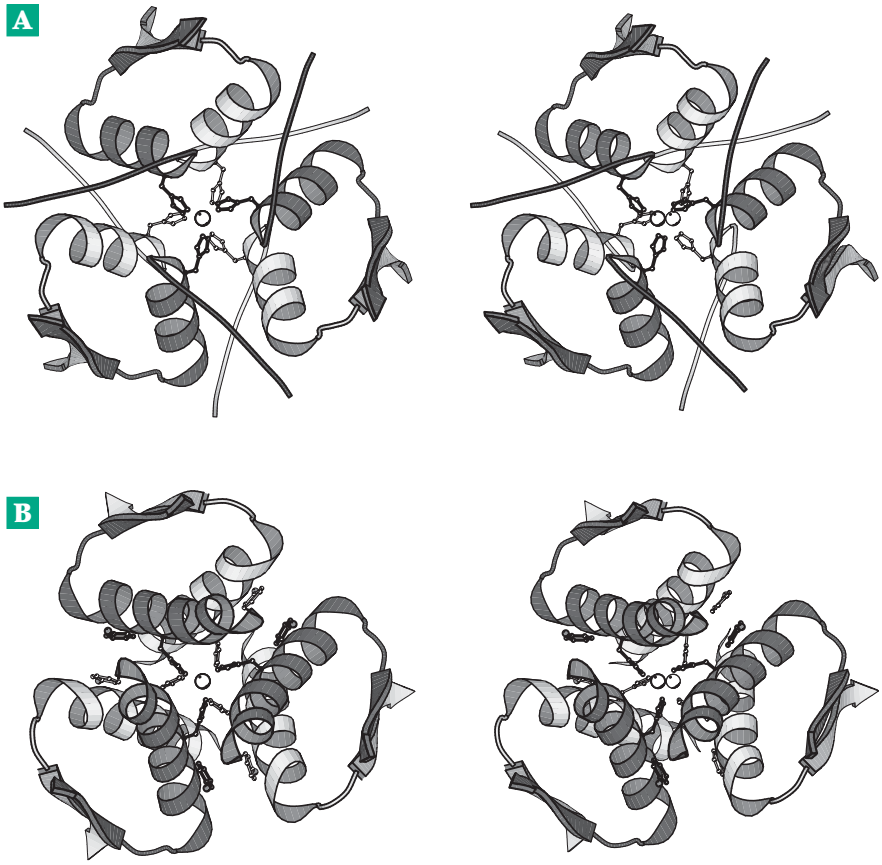



**Figure 7-18** Stereoscopic MolScript ribbon drawings of the B chains (A chains omitted) of (A) hexameric 2-zinc pig insulin. (B) A phenol complex of the same protein. Within each dimer the B chains are shaded differently. The $Zn^{2+}$ ions are represented by white spheres and the coordinating histidine side chains are shown. Six noncovalently bound phenol molecules can be seen, as can several conformational differences. From Whittingham *et al*.[87] Courtesy of Peter C. E. Moody.

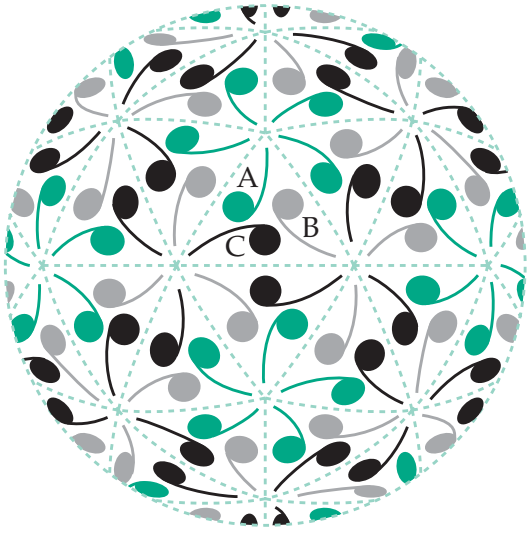



**Figure 7-19** Schematic icosahedrally symmetric structure with 180 subunits. The quasi-equivalent units A, B, and C are necessarily somewhat differently positioned with respect to their neighbors and must therefore assume different conformations in order to fit together tightly. From Harrison.[68]

to be violated.[98,99] Flexible arms tie the pentamers together.

Quasi-equivalence of subunits also provides the supercoil in bacterial flagella (Chapter 19) and accounts for some interesting aspects of the structure of tobacco mosaic virus. The protein subunits of the virus can exist either as a helix with 16.3 subunits per turn (Fig. 7-8) or as a flat ring of 17 subunits.[100] A very small conformational difference is involved. These rings dimerize but do not form larger aggregates. What is surprising is that the dimeric rings do not have dihedral symmetry, all of the subunits in the dimeric disk being oriented in one direction but with two different conformations. The disk may serve as an intermediate in virus assembly. The inner portions of the quasi-equivalent disk subunits have a jawlike appearance as if awaiting the incorporation of RNA. As the RNA becomes bound, the disks could dislocate to a "lock-washer" conformation to initiate and to propagate growth of the helical virus particle.[100,44a] However, there is uncertainty about this interpretation.[45,101]

Some enzymes, such as yeast hexokinase and creatine kinase (Chapter 12), associate in extremely asymmetric ways.[102] A dimer is formed by means of heterologous interactions but steric hindrance prevents the unsatisfied sets of interacting groups from joining with additional monomers to form higher polymers. As Galloway pointed out, many biological structures are not completely ordered but nevertheless possess well-defined and functionally important local relationships.[103]

***Regulatory subunits and multienzyme complexes.*** Proteins are often organized into large complexes, sometimes for the purpose of regulating metabolism. An example is **aspartate carbamoyltransferase** which catalyzes the first step in the synthesis of the pyrimidine rings of DNA and RNA (Chapter 25). The 310-kDa enzyme from *E. coli* can be dissociated into two 100-kDa trimers, referred to as **catalytic subunits**, and three 34-kDa dimers, the **regulatory subunits** which alter their conformations in response to changes in the ATP, UTP, and CTP concentrations.[104–107] The molecule is roughly triangular in shape[47,108] with a thickness of 9.2 nm and a length of the triangular side of 10.5 nm (Fig. 7-20). The symmetry is 3:2, i.e., it is dihedral with one 3-fold axis of rotation and three 2-fold axes. The two trimers of catalytic subunits lie face-to-face with the dimeric regulatory subunits fitting between them into the grooves around the edges of the trimers (Fig. 7-20). The dimers are not aligned exactly parallel with the 3-fold axis, but to avoid eclipsing, the upper half of the array is rotated around the 3-fold axis with respect to the lower half. In the center is an aqueous cavity of dimensions ~2.5 x 5.0 x 5.0 nm. The active sites of the enzyme are inside this cavity which is reached through six ~1.5-nm opening around the sides.

Many other oligomeric enzymes and other complex assemblies of more than one kind of protein subunit are known. For example, the **2-oxoacid dehydrogenases** are huge 2000- to 4000-kDa complexes containing three different proteins with different enzymatic activities in a cubic array (Fig. 15-14). The filaments of striated muscle (Chapter 19), antibodies and complement of blood (Chapter 31), and the tailed bacteriophages (Box 7-C ) all have complex molecular architectures.

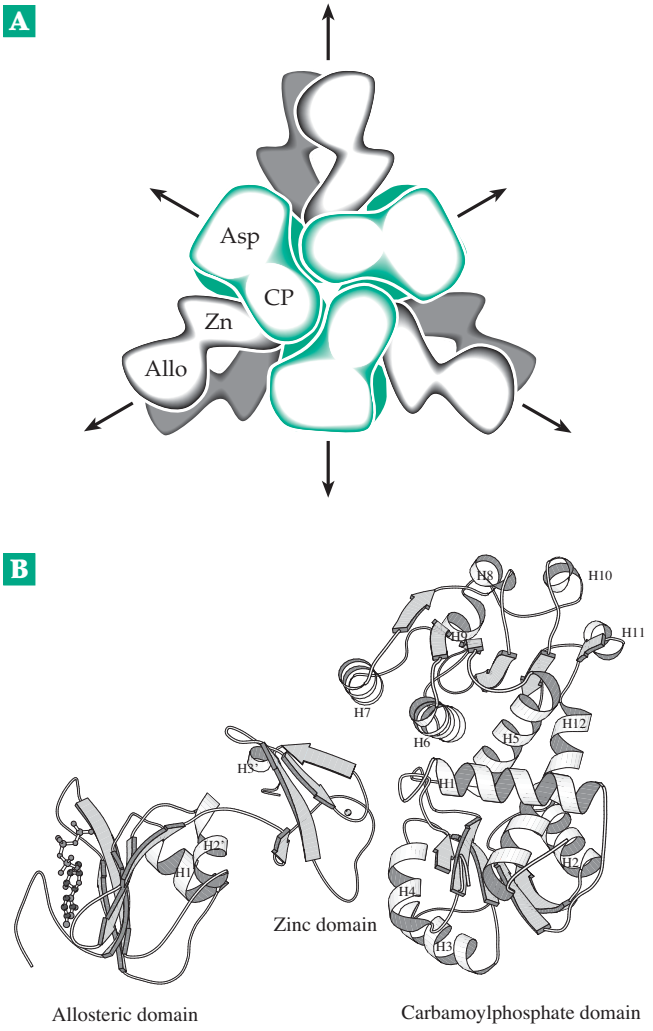


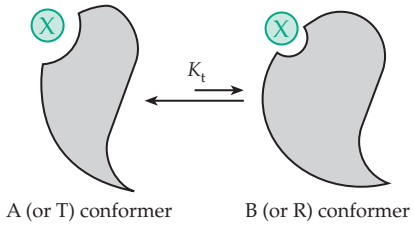

**Figure 7-20** (A) Subunit assembly of two $C_3$ catalytic trimers (green) and three $R_2$ regulatory dimers around the periphery in aspartate carbamoyltransferase. After Krause *et al.*[109] Courtesy of William N. Lipscomb. The aspartate- and carbamoylphosphate-binding domains of the catalytic subunits are labeled Asp and CP, respectivley, while the zinc and allosteric domains of the regulatory subunits are labeled Allo and Zn, respectively. (B) Ribbon drawing of a single pair of regulatory (left) and catalytic (right) subunits with the structural domains labeled. MolScript drawing from Thomas *et al.*[110]

## C. Cooperative Changes in Conformation

A substrate will bind better to some conformations of a protein than it will to others. This simple fact, together with the tendency for protein monomers to associate into clusters, allows for cooperative changes in conformation within oligomeric proteins. These changes provide the basis for important aspects of the regulation of enzymes and of metabolism. They impart cooperativity to the binding of small molecules such as that of oxygen to hemoglobin and of substrates and regulating molecules to enzymes. Many of the most fundamental and seemingly mysterious properties of living things are linked directly to cooperative changes within the fibrils, membranes, and other structures of the cell.

In 1965 a simple, appealing mathematical description of cooperative phenomena was suggested by Monod, Wyman, and Changeux[30,110a,110b] and focused new attention on the phenomenon. They suggested that conformational changes in protein subunits, which could be associated with altered binding characteristics, occur cooperatively within an oligomer. For example, binding of phenol to hexameric 2-zinc insulin (Fig. 7-18) could induce all six individual subunits to change their conformation together, preserving the $D_6$ symmetry of the complex. (In fact, it is more complex than this.[111]) The four subunits of hemoglobin could likewise change their conformation and affinity for $O_2$ synchronously. This is very nearly true and is of major physiological significance.

Consider an equilibrium (Eqs. 7-27 and 7-28) between protein molecules in two different conformations A and B (T and R in the MWC terminology) and containing a single binding site for molecule X. In the Monod–Wyman–Changeux (MWC) model the conformations are designated T (tense) and R (relaxed) but in the interest of providing a more general treatment the terminology used in this book is that of Koshland *et al*.[13,112–115]




(7-27)

$$K_t = [B] / [A] \tag{7-28}$$

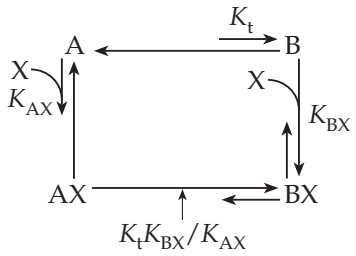
If the equilibrium constant $K_t$ is approximately 1, the two **conformers** have equal energies, but if $K_t < 1$, A is more stable than B.

### 1. Unequal Binding of Substrate and "Induced Fit"

Assume that conformer B binds X more strongly than does conformer A (as is suggested by the shapes of the binding sites in Eq. 7-27). The intrinsic binding constants to the A and B conformers $K_{AX}$ and $K_{BX}$ (or $K_T$ and $K_R$) are defined by Eq. 7-29:

$$K_{AX} = [AX] / [A][X]$$
$$K_{BX} = [BX] / [B][X]) \tag{7-29}$$

The entire set of equilibria for this system are shown in Eq. 7-30. Note that the constant relating



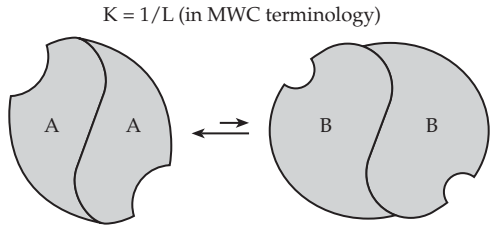

(7-30)

BX to AX is not independent of the other three constants but is given by the expression $K_tK_{BX} / K_{AX}$. Now consider the following situation. Suppose that A predominates in the absence of X but that X binds more tightly to B than to A. There will be largely either free A or BX in the equilibrium mixture with smaller amounts of AX and B. An interesting kinetic question arises. By which of the two possible pathways from A to BX (Eq. 7-30) will the reaction take place? The first possibility, assumed in the MWC model, is that X binds only to preformed B, which is present in a small amount in equilibrium with A. The second possibility is that X can bind to A but that AX is then rapidly converted to BX. We could say that *X induces a conformational change* that leads to a better fit. This is the basis for the **induced fit** theory of Koshland. Bear in mind that the equilibrium constants can give us the equilibrium concentrations of all four forms in Eq. 7-30. However, rates of reaction are often important in metabolism and we cannot say *a priori* which of the two pathways will be followed.

If $K_{BX} / K_{AX}$ is very large, an insignificant amount of AX will be present at equilibrium. In such a case there is no way experimentally to determine $K_{AX}$. The two constants $K_t$ and $K_{BX}$ are sufficient to describe the *equilibria* but an induced fit mechanism may still hold.

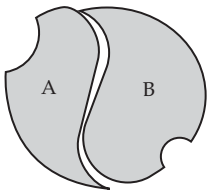
Now consider the association of A and B to form oligomers in which the intrinsic binding constants $K_{AX}$ and $K_{BX}$ have the same values as in the monomers. Since more enzymes apparently exist as isologous dimers than as any other oligomeric form,[116] it is appropriate to consider the behavior of such dimers in some detail. Monod *et al*. emphasized that both

conformers A and B (T and R) can associate to form isologous dimers in which symmetry is preserved (Eq. 7-31).




(7-31)

On the other hand, association of B and A would lead to an unsymmetric dimer in which bonding between subunits might be poor:




Mixed AB dimer
which associates weakly

In the MWC treatment, the assumption is made that the mixed dimer AB can be neglected entirely. However, a more general treatment requires that we consider all dimeric forms. The formation constants $K_{AA}$, $K_{BB}$, and $K_{AB}$ are defined as follows[13,113] (Eqs. 7-32 to 7-34; note the statistical factor of 2 relating $K_{AB}$ to the association constant $K_f$):

$$2A \rightleftarrows A_2 \qquad K_{AA} = \frac{[A_2]}{[A]^2} \tag{7-32}$$

$$2B \rightleftarrows B_2 \qquad K_{BB} = \frac{[B_2]}{[B]^2} = \frac{[B_2]}{K_t^2[A]^2} \tag{7-33}$$

$$A + B \rightleftarrows AB \qquad K_f = 2K_{AB} = \frac{[AB]}{[A][B]} = \frac{[AB]}{K_t[A]^2} \tag{7-34}$$

## 2. Binding Equilibria for a Dimerizing Protein

All of the equilibria of Eqs. 7-28 through 7-34 involved in formation of dimers $A_2$, AB, and $B_2$ and in the binding of one or two molecules of X per dimer are depicted in Fig. 7-21. Above each arrow the microscopic constant associated with that step is shown multiplied by an appropriate statistical factor. The fractional saturation Y is given by Eq. 7-35. Each of the nine terms in the numerator gives the concentration of

$$2Y \text{ (based on dimer)} = \frac{\begin{array}{r}[AX] + [BX] + [A_2X] + 2[A_2X_2] + [ABX] \\ + [AXB] + 2[ABX_2] + [B_2X] + 2[B_2X_2]\end{array}}{\begin{array}{r}\tfrac{1}{2}([A] + [AX] + [B] + [BX]) + [A_2] + [A_2X] \\ + [A_2X_2] + [AB] + [ABX] + [AXB] \\ + [ABX_2] + [B_2] + [B_2X] + [B_2X_2]\end{array}} \tag{7-35}$$

bound X represented by one of the nine forms containing X in Fig. 7-21. The 14 terms in the denominator represent the concentration of protein in each form including those containing no bound X. Protein concentrations are given in terms of the molecular mass of the dimer; hence, some of the terms in the denominator are multiplied by 1/2.

All of the terms in both the numerator and the denominator of Eq. 7-35 can be related back to [X], using the microscopic constants from Fig. 7-21 to give an equation (comparable to Eq. 7-8) which presents Y in terms of [X], $K_{AX}$ and $K_{BX}$, $K_t$, and the interaction constants $K_{AA}$, $K_{AB}$, and $K_{BB}$. Since the equation is too complex to grasp immediately, let us consider several specific cases in which it can be simplified.

***The Monod-Wyman-Changeux (MWC) model.*** If both $K_{AA}$ and $K_{BB}$ are large enough, there will be no dissociation into monomers. The transition between conformation A and conformation B can occur cooperatively within the dimer or higher oligomer, and the mathematical relationships shown in Fig. 7-21 are still appropriate. One further restriction is needed to describe the MWC model. Only symmetric dimers are allowed. That is, $K_{AA}$ and $K_{BB} >> K_{AB}$ (see Eq. 7-31), and only those equilibria indicated with green arrows in Fig. 7-21 need be considered.[30] In the absence of ligand X, the ratio $[B_2]/[A_2]$ is a constant, 1/L in the MWC terminology (Eq. 7-36; see also Eq. 7-31).

$$\frac{[B_2]}{[A_2]} = \frac{1}{L} = \frac{K_{BB}}{K_{AA}} K_t^2 \tag{7-36}$$

Both of the association constants $K_{AA}$ and $K_{BB}$ and the transformation constant $K_t$ affect the position of the equilibrium. Thus, a low ratio of $[B_2]$ to $[A_2]$ could result if $K_{BB}$ and $K_{AA}$ were similar but $K_t$ was small. If $K_t$ were ~1 a low ratio could still arise because $K_{AA} >$

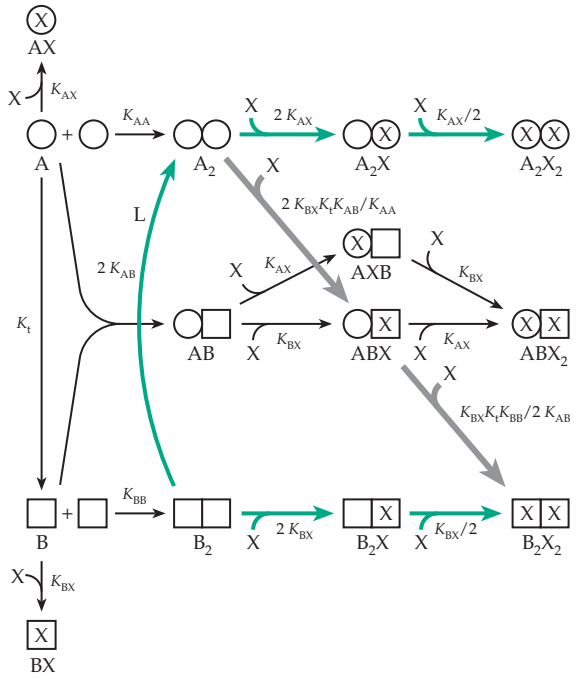



**Figure 7-21** Possible forms of a dimerizing protein existing in two conformations with a single binding site per protomer for X. Green arrows indicate equilibria considered by MWC. Solid arrows indicate equilibria considered by Koshland *et al.*[13,114] Heavy gray arrows are for the simplest induced fit model with no dissociation of the dimer. Note that all equilibria are regarded as reversible (despite the unidirectional arrows). $K_{AX}$ and $K_{BX}$ are assumed the same for subunits in monomeric and dimeric forms.

$K_{BB}$, i.e., because the subunits are associated more tightly in $A_2$ than in $B_2$. For this case Eq. 7-35 simplifies to Eq. 7-37.

$$2Y = \frac{[A_2X] + 2[A_2X_2] + [B_2X] + 2[B_2X_2]}{[A_2] + [A_2X] + [A_2X_2] + [B_2] + [B_2X] + [B_2X_2]}$$

$$= \frac{\begin{array}{c}2K_{AA}K_{AX}[X] + 2K_{AA}K_{AX}^2[X]^2 \\ + 2K_{BB}K_{BX}K_t^2[X] + 2K_{BB}K_{BX}K_t^2[X]^2\end{array}}{\begin{array}{c}K_{AA} + 2K_{AA}K_{AX}[X] + K_{AA}K_{AX}^2[X]^2 + K_{BB}K_t^2 \\ + 2K_{BB}K_{BX}K_t^2[X] + K_{BB}K_{BX}^2K_t^2[X]^2\end{array}} \qquad (7\text{-}37)$$

Substituting from Eq. 7-36 into Eq. 7-37 we obtain (Eq. 7-38):

$$Y \text{ (for dimer)} = \frac{L \cdot K_{AX}[X](1 + K_{AX}[X]) + K_{BX}[X](1 + K_{BX}[X])}{L(1 + K_{AX}[X])^2 + (1 + K_{BX}[X])^2} \qquad (7\text{-}38)$$

For an oligomer with $n$ subunits Monod *et al.* assumed that all sites in either conformer are independent and equivalent. The equation for Y (based on Eq. 7-17) is

$$Y = \frac{L \cdot K_{AX}[X](1 + K_{AX}[X])^{n-1} + K_{BX}[X](1 + K_{BX}[X])^{n-1}}{L(1 + K_{AX}[X])^n + (1 + K_{BX}[X])^n} \qquad (7\text{-}39)$$

We assume initially that $B_2$ binds X more strongly than does $A_2$. Hence, if the equilibrium in Eq. 7-36 favors $B_2$ strongly ($L$ is small), the addition of X to the system will not shift the equilibrium between the two conformations and binding will be noncooperative (Eq. 7-39 will reduce to Eq. 7-18). However, if the equilibrium favors $A_2$ ($L$ is large), addition of X will shift the equilibrium in favor of $B_2$ (which binds X more tightly). Furthermore, since the expression for Y (Eq. 7-39) contains a term in $K^2_{BX}[X]^2$ in the numerator, binding will tend to be cooperative. In the extreme case that $L$ is large and $K_{AX} \sim 0$, most of the terms in Eq. 7-39 drop out and it approaches the equation previously given for completely cooperative binding (Eq. 7-21) with $K = K_{BX}^2L$. With other values of $K_{AX}$, $K_{BX}$, and $L$ incomplete cooperativity is observed.[112]

Further development of the MWC theory as it applies to enzyme kinetics is given in Chapter 9, Section B.

***The induced fit model.*** In this model, only $A_2$, ABX, and $B_2X_2$ are considered (heavy arrows in Fig. 7-15).[13,114] The expression for 2Y is:

$$2Y = \frac{[ABX] + 2[B_2X_2]}{[A_2] + [ABX] + [B_2X_2]}$$

$$= \frac{2K_{BX}K_t\frac{K_{AB}}{K_{AA}}[X] + 2(K_{BX}K_t)^2\frac{K_{BB}}{K_{AA}}[X]^2}{1 + 2K_{BX}K_t\frac{K_{AB}}{K_{AA}}[X] + (K_{BX}K_t)^2\frac{K_{BB}}{K_{AA}}[X]^2} \qquad (7\text{-}40)$$

The constants used here are defined by Eqs. 7-8 through 7-10 and differ from those of Koshland, who sometimes arbitrarily set $K_{AA} = 1$ and redefined $K_{BB}$ as an *interaction constant* equal to $K_{BB} / K_{AA}$. Although this simplifies the algebra it is appropriate only for completely associated systems and might prove confusing.

When $K_{AB}$ is small (no "mixed" dimer) Eq. 7-16 also simplifies to Eq. 7-45 for completely cooperative binding with the value $K$ given by Eq. 7-17. On the other hand, if $K_{AB}$ is large compared to $K_{AA}$ and $K_{BB}$, anticooperativity (negative cooperativity) will be observed. The saturation curve will contain two separate steps just as in the binding of protons by succinate dianion (Fig. 7-5).

$$\bar{K} = K_{BX}{}^2 K_t{}^2 \frac{K_{BB}}{K_{AA}} = \frac{K_{BX}{}^2}{L} \qquad (7\text{-}41)$$

***One conformational state dissociated.*** It may happen that $A_2$ is a dimer but that $B_2$ dissociates into monomers because $K_{BB}$ is very small. In such a case binding of X leads to dissociation of the protein. A well-known example is provided by hemoglobin of the lamprey which is a dimer and which dissociates to a monomer upon binding of oxygen.[117] Equation 7-11 simplifies to Eq. 7-42. The reader may wish to consider whether the weakly cooperative binding of oxygen by lamprey hemoglobin is predicted by this equation.

$$2Y = \frac{[BX]}{[A_2] + \frac{1}{2}[BX]} \qquad (7\text{-}42)$$

Look again at the expression for $L$, the constant determining the relative amounts of a protein in conformations A and B in the absence of ligand. From Eq. 7-36 we see that a large value of $L$ (conformer A favored) can result either because $K_t$ is very small or because $K_{BB} << K_{AA}$. Thus, if $K_t \sim 1$ and $L$ is large, the subunits must associate much more weakly in $B_2$ than in $A_2$ and the chances are that binding of X will dissociate the molecule as in the case of lamprey hemoglobin. On the other hand, if $K_t$ is very small, implying that the molecule is held in conformation A because of some intrinsically more stable folding pattern in that conformation, $K_{BB}$ might exceed $K_{AA}$ very much; if $K_{AA}$ were low enough $A_2$ could be completely dissociated. Binding of ligand would lead to association and to cooperative binding. This can be verified by writing down the appropriate terms from Eq. 7-35.

## 3. Higher Oligomers

Mathematical treatment of binding curves for oligomers containing more than two subunits is complex, but the algebra is straightforward. A computer can be programmed to do necessary calculations. Avoid picking an equation from the literature and assuming that it will be satisfactory. Consider the two tetrameric structures in Fig. 7-22. In the isologous square (also shown in Fig. 7-11) separate contributions to the free energy of binding can be assigned to the individual pairs of interactions *aj* and *bk*.

Thus, following Cornish-Bowden and Koshland[114] for assembly of the tetramer (Eq. 7-43):

$$\Delta G_f = 2\,\Delta G_{ajAA} + 2\,\Delta G_{bkAA}$$
$$K_f = K_{ajAA}^2 K_{bkAA}^2 \qquad (7\text{-}43)$$

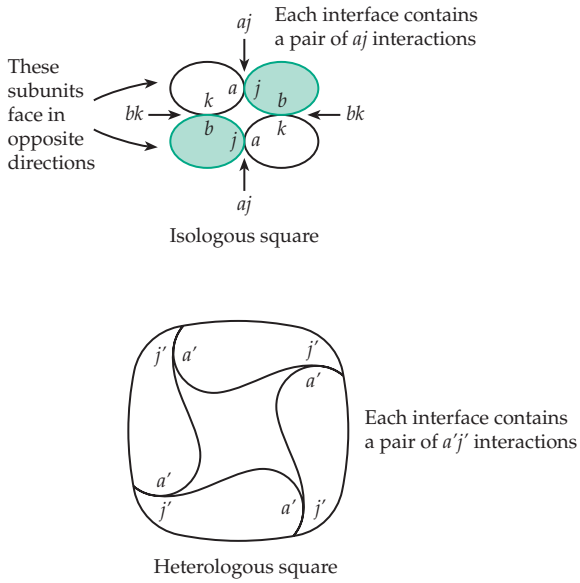
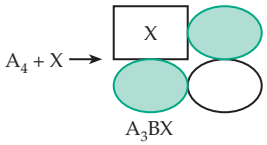



**Figure 7-22** Comparison of the interactions in isologous (dihedral) and heterologous (cyclic) square configurations of subunits.

Since Gibbs energies are additive, the formation constant will be the product of formation constants representing the individual interactions; thus, $K_{ajAA}$ represents the formation constant of a dimer in which only the *aj* pair of bonds is formed.

In the isologous tetrahedron (Fig. 7-5) the third set of paired interactions *cl* must be taken into account. (However, the third interaction constant will not be an association constant of the type represented by $K_{ajAA}$ and $K_{bkAA}$ but a dimensionless number.) On the other hand, the heterologous square has only a single interaction constant.

Now consider the binding of one molecule of X to the isologous tetramer with a conformational change in one subunit (Eq. 7-44). We see that one pair of *aj* interactions and one pair of *bk* interactions have been

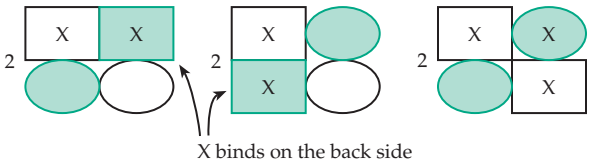



(7-44)

altered. The equilibrium constant for the binding of X to the tetramer will be (Eq. 7-45) in which the 4 is a

$$K = 4 \frac{K_{ajAB} K_{bkAB}}{K_{ajAA} K_{bkAA}} K_{BX} K_t \qquad (7\text{-}45)$$

statistical factor arising from the fact that there are four different ways in which to form $A_3BX$. When a second molecule of X is added three geometrical arrangements are possible:




Each one can be formed in two ways. It is a simple matter to write down the microscopic constants for addition of the second molecule of X as the sum of three terms. Because values of the constants for *aj* and *bk* differ, it will be clear that the three ways of adding the second molecule of X are not equally probable. Thus, the oligomer will show preferred orders of "loading" with ligand X.

Two different geometries for the heterologous tetramer are possible in form $A_2B_2X_2$. Again, the different arrangements need not be equally probable and the relative distribution of each will be determined by the specific values of the interaction constants. In the heterologous tetramer $A_4$ only one type of an interaction is present between subunits. However, as soon as a single molecule of X is bound and one subunit of conformation B is present, two kinds of *aj* interactions exist. (One in which group *a* is present in conformation A and the other in which it is present in conformation B.) Since these interactions always occur in equal numbers they can be lumped together.

While the foregoing may seem like an unnecessarily long exercise, it should provide a basic approach which can be applied to specific problems. However, remember that mathematical models require simplification. Real proteins often have more than two stable conformations.[117] The entire outside surface of a protein is made up of potential binding sites for a number of different molecules, both small and large. Filling of almost any of these sites can affect the functioning of a protein.

## D. The Oxygen-Carrying Proteins

### 1. Myoglobin and Hemoglobin

The most studied example of a conformational change in a multisubunit protein induced by binding of a small molecule is provided by the cooperative binding of oxygen to hemoglobin.[118–120] Mammalian hemoglobin is an $\alpha_2\beta_2$ tetramer of ~16-kDa subunits, each containing 140 – 150 residues. Within each subunit the peptide chain folds in a characteristic largely α-helical pattern around a single large flat iron-containing ring structure called **heme** (Fig. 7-23). The folding is essentially the same in all hemoglobins, both in the α and β subunits and in the monomeric muscle oxygen storage protein, myoglobin. Amino acid residues are customerily designated by their position in one of the eight helices A–H. The imidazole group of histidine F-8 is coordinated with the iron in the center of the heme on the "proximal" side. The other side of the iron atom (the "distal" side) is the site of binding of a single molecule of $O_2$.

Although the folding of the peptide chain is almost the same in both subunits, and almost identical to that of myoglobin,[118,121,122] there are numerous differences in the amino acid sequence. If it were not for these differences, hemoglobin would be a highly symmetric molecule with the bonding pattern indicated in Fig. 7-5 with three 2-fold axes of rotation. In fact, hemoglobin has one true axis of rotation and two pseudo-twofold axes. There are two sets of true isologous interactions (those between the two α subunits and between the two β subunits) and two pairs of unsymmetrical interactions (between α and β subunits). The nearly symmetric orientation of different portions of the peptide backbone is clearly seen in the beautiful drawings of Geis.[119]

The contact region involved in one pair of interactions in hemoglobin ($\alpha_1\beta_1$) is more extensive than the other. There is close contact between 34 different amino acid side chains and 110 atoms lie within 0.4 nm of each other.[118] Hydrophobic bonding is the principal force holding the two subunits together, and only a few reciprocal contacts of the type found in a true isologous bond remain. The second contact designated $\alpha_1\beta_2$ involves only 19 residues and a total of 80 atoms. Because this interaction is weaker, hemoglobin dissociates relatively easily into αβ dimers held together by the $\alpha_1\beta_1$ contacts and motion occurs along the $\alpha_1\beta_2$ contacts during oxygenation. The truly isologous interactions (i.e., αα and ββ ) are weak because the identical protomers hardly touch each other.

***The binding of oxygen.*** Curves of percentage oxygenation (Y) vs the partial pressure of $O_2$ are given in Fig. 7-24 and illustrate the high degree of cooperativity. Depending upon conditions, values of $n_{Hill}$ (Eq. 7-24) may be as high as 3. As a result of this cooperativity the hemoglobin, in the capillaries of the lungs at a partial pressure of oxygen of ~100 mm of mercury, is nearly saturated with oxygen. However, when the red cells pass through the capillaries of tissues in which oxygen is utilized the partial pressure of oxygen falls to about 5 mm of mercury. The cooperativity means that the oxygen is more completely "unloaded" than it would be if all four heme groups acted independently.

Deoxyhemoglobin has a low affinity for $O_2$, but the observed cooperativity in binding implies that in the fully oxygenated state the $O_2$ is held with a high

affinity. The monomeric myoglobin also has a high affinity for oxygen, as does the abnormal **hemoglobin H**, which is made up of four β subunits. The latter also completely lacks cooperativity in binding.[123] These results can be interpreted according to the MWC model to indicate that deoxyhemoglobin exists in the T (A) conformation, whereas oxyhemoglobin is in the R (B) conformation. Myoglobin stays in the R conformation in *both* states of oxygenation as do the separated α and β chains of hemoglobin. The subunits of hemoglobin H also appear to be frozen in the R conformation, even though the quaternary structure is similar to that of deoxyhemoglobin.[123,124]

Oxygenation curves of hemoglobin are often fitted with the Adair equation (Eq. 7-12). Thus, at pH 7.4 under the conditions given in Table 7-2, Imai[125] found for the successive formation constants $K_1 = 0.004$, $K_2 = 0.009$, $K_3 = 0.002$ and $K_4 = 0.95$ in units of mm $Hg^{-1}$. From the definition of a formation constant the oxygen pressure $P_{O_2}$ required for 50% oxygenation in the first step will be at $P_{O_2} = 1/K_f$ or $\log P_{O_2} = \log (1/.004) = 2.4$. This is a high oxygen pressure, far to the right side of the oxygenation curve in Fig. 7-24A. However, $\log K_4$ is about 0.02, well to the left on the oxygenation curve. From these formation constants we can say that after three of the subunits have become oxygenated the affinity of the remaining subunit has increased about 300-fold when the concentration of the effector 2,3-bisphosphoglycerate is present at the normal physiological concentraton (see Section 4).[125–127]

However, we must ask what uncertainties are present in the data used to obtain these constants. To extract four successive binding constants from a curve like that in Fig. 7-24A is extremely difficult.[129,130] This fact has encouraged the widespread use of the simpler MWC model.[30a,127a] When the same data were treated by Imai[131] using the MWC model it was found that $L = 2.8 \times 10^6$ and $c = K_{f(T)} / K_{f(R)} = 0.0038$. Changes in

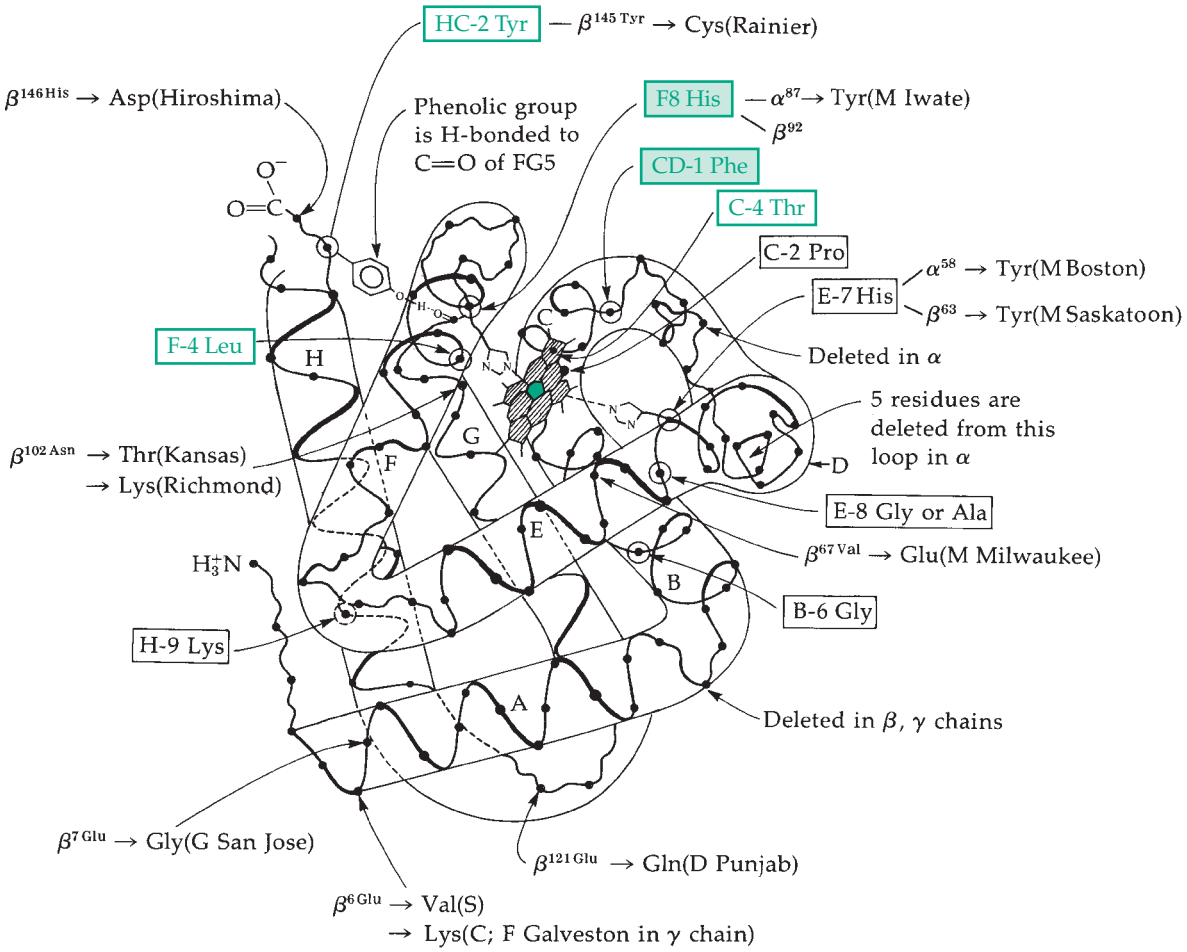



**Figure 7-23** Folding pattern of the hemoglobin monomers. The pattern shown is for the β chain of human hemoglobin. Some of the differences between this and the α chain and myoglobin are indicated. Evolutionarily conserved residues are indicated by boxes, ▭ highly conserved, ▭ invariant. Other markings show substitutions observed in some abnormal human hemoglobins. Conserved residues are numbered according to their location in one of the helices A–H, while mutant hemoglobins are indicated by the position of the substitution in the entire α and β chain.

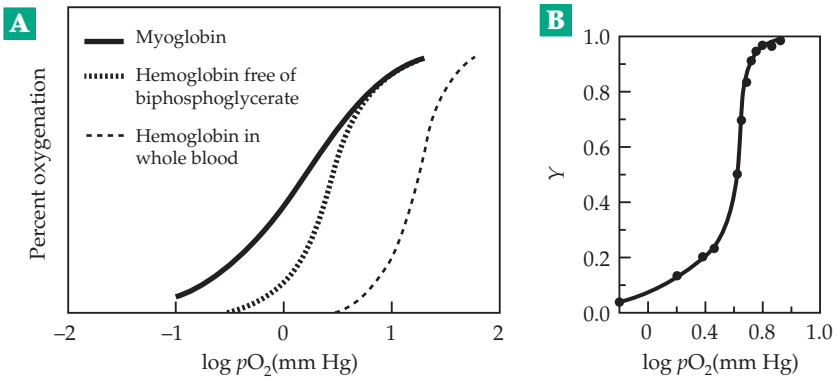



**Figure 7-24** Cooperative binding of oxygen by hemoglobins. (A) Binding curve for myoglobin (noncooperative) and for hemoglobin in the absence and presence (in whole blood) of 2,3-bisphosphoglycerate. Oxygen affinity is decreased by bisphosphoglycerate. (B) Saturation curve for hemoglobin (erythrocruorin) of *Arenicola*, a spiny annelid worm. The molecule contains 192 subunits and 96 hemes. It shows very strong cooperativity with $n_{Hill} \sim 6$. From (A) Benesch and Benesch,[123] and (B) Waxman.[128]

enthalpy, entropy, and Gibbs energy are given in Table 7-2. Kinetic data[131a] as well as $O_2$-binding measurements with single crystals[132] are partially consistent with the MWC model.[133] However, the discovery of a third quaternary structure of hemoglobin, similar to the R state but distinct from it,[134–136a] emphasizes the complexity of this allosteric molecule.

Hemoglobin tetramers tend to stay tightly associated but some dissociation of oxyhemoglobin into dimers does occur ($K_f = 7 \times 10^5$ M).[137,138] Deoxyhemoglobin is about 40,000 times more tightly associated. All of the equilibria involved are strongly affected by pH and by the presence of salts such as NaCl.[126,139] This is in part a reflection of the strong role of ionic interactions in holding together the subunits in the T state as is discussed in the following sections.[127]

**TABLE 7-2**
**Thermodynamic Functions for Oxygenation of Hemoglobin[a]**

| Reaction | $\Delta H$ (kJ mol$^{-1}$) | $\Delta S$ (J°K$^{-1}$ mol$^{-1}$) | $\Delta G$ (kJ mol$^{-1}$) | $K_f$ |
|---|---|---|---|---|
| $T \rightarrow T(O_2)_4$ | −51 ± 1 | −154 ± 4 | −5.0 ± 1.7 | 7.5 |
| $R \rightarrow R(O_2)_4$ | −62 ± 2 | −146 ± 6 | −19 ± 2 | $2.0 \times 10^3$ |
| $T \rightarrow R$ (unoxygenated) | −70 ± 7 | −111 ± 25 | −37 ± 10 | $3.6 \times 10^{-7}$ |
| $T(O_2)_4 \rightarrow R(O_2)_4$ (oxygenated) | | | −19 | $2.1 \times 10^3$ |

Parameters for MWC model

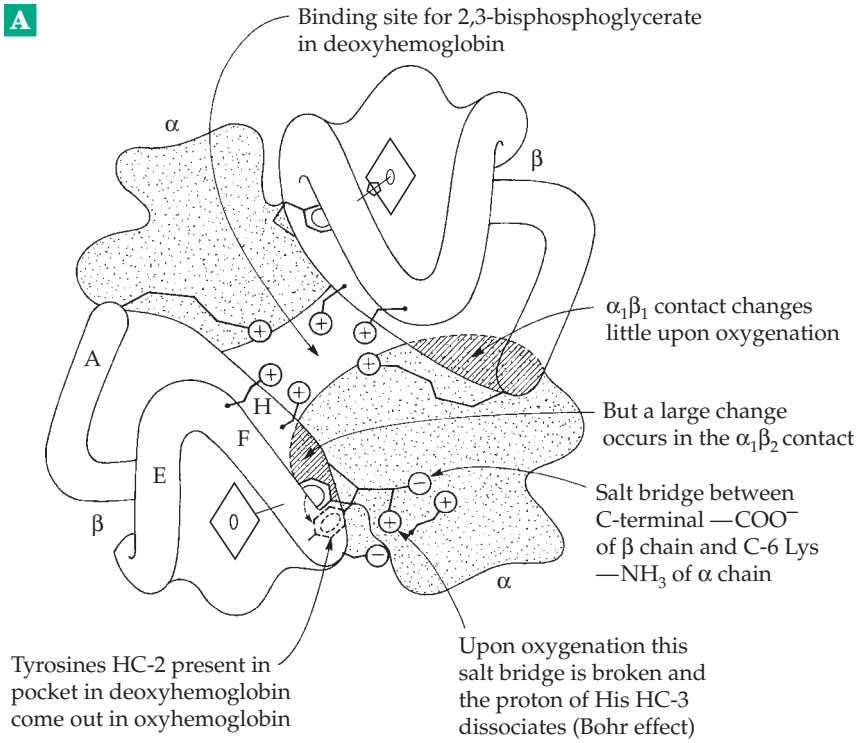
$L = (3.6 \times 10^{-7})^{-1} = 2.8 \times 10^6$

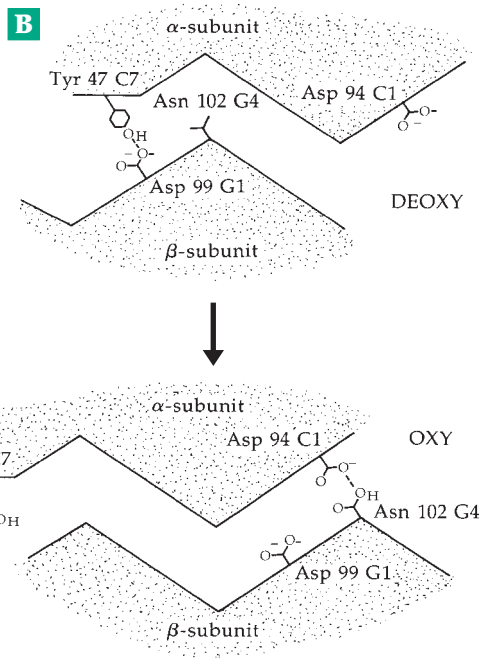
$c = K_{f(T)} / K_{f(R)} = 0.0038$

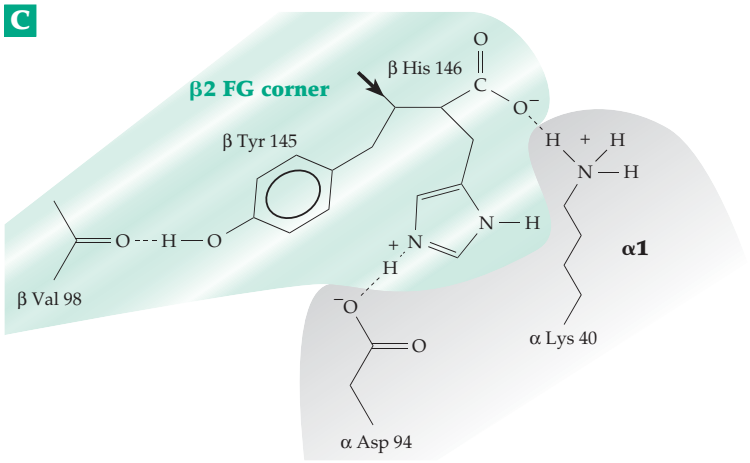
[a] From Imai, K. (1979) *J. Mol. Biol.* **133**, 233–247. The measurements were made at pH 7.4 in the presence of 0.1 M chloride ion and 2 mM 2,3-bisphosphoglycerate to mimic physiological conditions. The values of $\Delta H$, $\Delta S$, and $\Delta G$ given are per mole of heme, i.e., per monomer unit. They must be multiplied by 4 to correspond to the reactions as shown for the tetramer.

***Structural changes accompanying oxygen binding.*** Perutz and associates, using X-ray crystallography, found *small* but real differences in the conformation of the subunits of deoxy- and oxy- hemoglobin.[140,141] More striking is the fact that upon oxygenation, both α and β subunits undergo substantial amounts of *rotation*, the net result being that the hemes of the two β subunits move about 0.07 nm closer together in the oxy form than in the deoxy form. Within the $\alpha_1\beta_1$ contacts (Fig. 7-25) little change is seen. On the other hand, contact $\alpha_1\beta_2$, the "allosteric interface," is altered drastically. As Perutz expressed it, there is a "jump in the dovetailing" of the CD region of the subunit relative to the FG region of the β subunit. The hydrogen-bonding pattern is also changed. A major difference is seen in the hydrogen-bonded salt bridges present at the ends of the molecules of deoxyhemoglobin. The $-NH_3^+$ group of Lys H-10 in each α subunit is hydrogen bonded to the carboxyl group of the C-terminal arginine of the opposite α chain. The guanidinium group of each C-terminal arginine is hydrogen bonded to the carboxyl group of Asp H-9 in the opposite α chain. It is also hydrogen bonded to an inorganic anion (phosphate or $Cl^-$), which in turn is hydrogen-bonded to the α amino group of Val 1 of the opposite α chain[142] forming a pair of isologous interactions.

At the other end of the molecule, the C-terminal group of His 146 of each β chain binds to the amino group of Lys C-6 of the α chain, while the imidazole side chain binds to Asp FG-1 of the same β chain (Fig. 7-25). These salt bridges appear to provide extra stability to



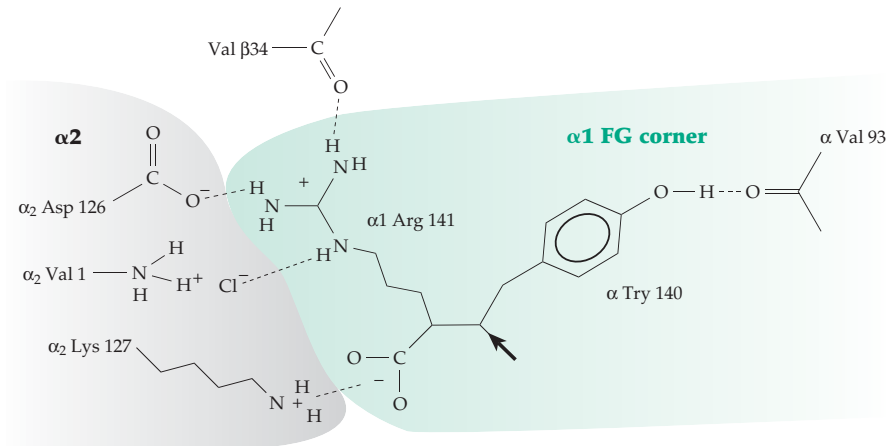




**Figure 7-25** (A) Structural changes occurring upon oxygenation of hemoglobin. After Dickerson[144] and Perutz.[143] (B) "Rotation at the contact $\alpha_1\beta_2$ causes a jump in the dovetailing of the CD region of α relative to the FG region of β and a switch of hydrogen bonds as shown".[143] (C) Some details of the salt bridges.

deoxyhemoglobin and account for the high value of the constant L. In deoxyhemoglobin the side chain of the highly conserved Tyr HC-2 lies tucked into a pocket between the H and F helices and is hydrogen bonded to the main chain carbonyl of residue FG-5 (Figs. 7-23 and 7-25). Upon oxygenation this tyrosine in each subunit is released from its pocket; the salt bridges at the ends of the molecules are broken and the subunit shifts into the new bonding pattern characteristic of oxyhemoglobin.[143,144] Cooperativity in $O_2$ binding is absent or greatly decreased in mutant hemoglobins with substitutions in the residues involved in these salt bridges[145–147] or in residues lying in the $\alpha_1\beta_2$ interface.[148,149]

How does the binding of $O_2$ to the iron of heme trigger the conformational change in hemoglobin? An enormous amount of effort by many people has been expended in trying to answer this question. As is pointed out in Chapter 16, the iron atom in deoxyhemoglobin lies a little outside the plane of the heme rings. When oxygenation occurs the iron atom moves toward the oxygen and into the plane of the heme.[150,151] This movement probably amounts to only about 0.05 nm. Nevertheless, this small displacement evidently induces the other structural changes that are observed. The iron pulls the side chain of histidine F-8 with it and moves helix F which is also hydrogen bonded to this imidazole ring. Because of the tight packing of the various groups this motion cannot occur freely but is accompanied by a movement of the F helix by 0.1 nm across the heme plane. These movements may induce additional structural changes in the irregularly folded FG corners that allow the subunits to shift to the new stable position of the R state. All four subunits appear to change conformation together. This conformational change must also cause the affinity for oxygen of any unoxygenated subunits to rise dramatically, presumably by shifting the iron atoms into the planes of the heme rings. This ensures the cooperative loading of the protein by $O_2$.

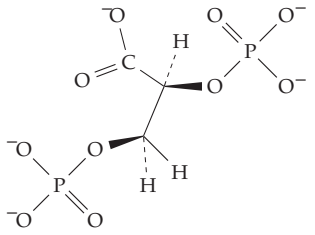
While there is no doubt that the iron atom moves upon oxygenation, it is not obvious that this will lead to the observed cooperativity. Oxygenated heme has some of the characteristics of an Fe(III)–peroxide anion complex.[152] The iron atom acquires an increased positive charge upon oxygenation by donating an electron for bond formation.

$$Fe^{2+}\ (\text{deoxy}) + O_2 \longrightarrow Fe^{3+}\text{–}\ O_2^{-} \qquad (7\text{-}46)$$

This may transmit an electronic effect through either His F-8 or the heme ring to the nearby $\alpha_1\beta_2$ interface and also affect the subunit interactions. The reaction of various heme proteins with oxygen is discussed further in Chapter 16.

***The Bohr effect and allosteric regulators.*** The breaking of the salt bridges at the ends of the hemoglobin molecule upon oxygenation has another important result. The $pK_a$ values of the N-terminal valines of the α subunits and of His HC-3 of the β subunits are abnormally high in the deoxy form because they are tied up in the salt bridges. In the oxy form in which the groups are free, the $pK_a$ values are lower. If hemoglobin is held at a constant pH of 7, these protons dissociate upon oxygenation. This **Bohr effect**, described in 1904,[153–156] is important because acidification of hemoglobin stabilizes the deoxy form. In capillaries in which oxygen pressure is low and in which carbon dioxide and lactic acid may have accumulated, the lowering of the pH causes oxyhemoglobin to release oxygen more efficiently. These effects are also strongly dependent on the presence of chloride ions.[139,142,143,157,158]

Just as the conformational equilibria in hemoglobin can be shifted by attachment of oxygen to the heme groups, so the binding of certain other molecules at different sites can also affect the conformation. Such compounds are called **allosteric effectors** or **regulators** because they bind at a site other than the "active site." They are considered in more detail in Chapter 9. An important allosteric effector for human hemoglobin is **2,3-bisphosphoglycerate**, a compound found in human red blood cells in a high concentration approximately equimolar with that of hemoglobin. One molecule of bisphosphoglycerate binds to a hemoglobin tetramer in the deoxy form with $K_f = 1.4 \times 10^5$ but has only half this affinity for oxyhemoglobin.[159] X-ray crystallography shows that bisphosphoglycerate binds between the two β chains of deoxyhemoglobin directly on the twofold axis (Fig. 7-26).[159] Because of the presence



2,3-Bisphosphoglycerate

of 2,3-bisphosphoglycerate in erythrocytes the affinity of oxygen for hemoglobin in whole blood is less than that for isolated hemoglobin[160,161] (Fig. 7-24). This is important because it allows a larger fraction of the oxygen carried to be unloaded from red corpuscles in body tissues. The bisphosphoglycerate level of red cells varies with physiological conditions, e.g., people living at higher elevations have a higher concentration.[161] It has been suggested that artificial manipulation of the level of this regulatory substance in erythrocytes may be of clinical usefulness for disorders in oxygen transport.

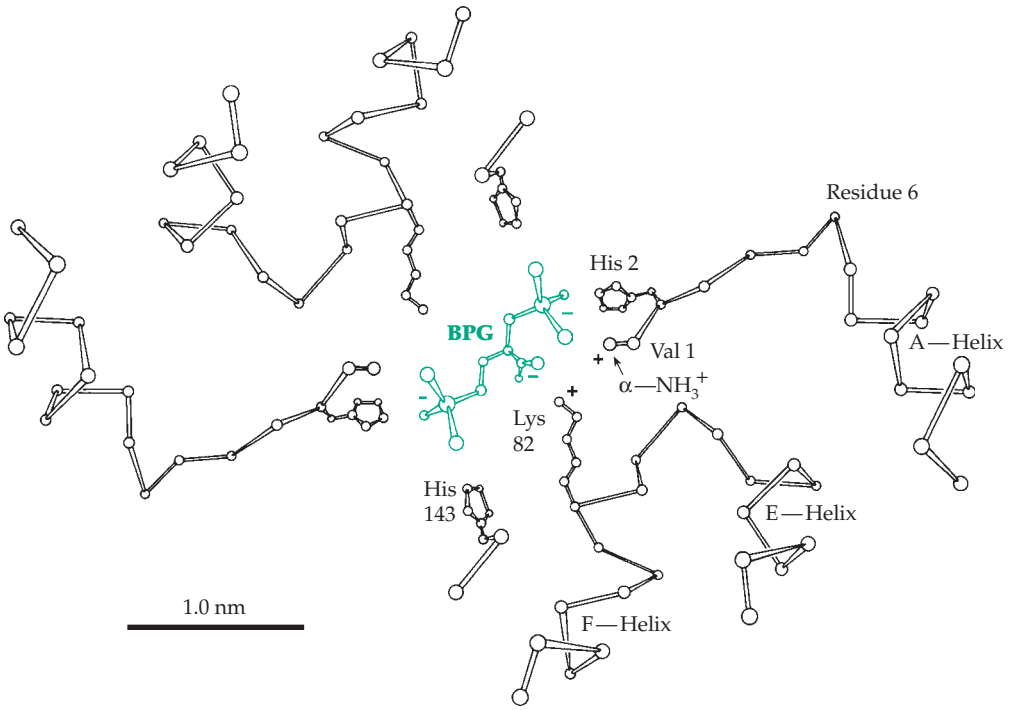
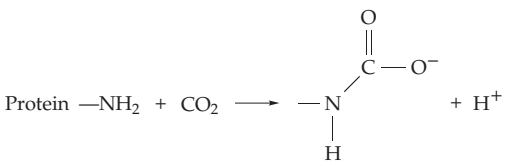



**Figure 7-26** The allosteric effects of 2,3-bisphosphoglycerate (BPG) bound to the β chains of human deoxyhemoglobin. The phosphate groups of the BPG form salt bridges with valines 1 and histidines 2 and 143 of both β chains and with lysine 82 of one chain. This binding pulls the A helix and residue 6 toward the E helix. From Arnone.[159]

Not all species contain 2,3-bisphosphoglycerate in their erythrocytes. In birds and turtles its function appears to be served by inositol pentaphosphate. In crocodiles the site between the two β chains that binds organic phosphates in other species has been modified so that it binds bicarbonate ion, $HCO_3^-$, specifically. This ion, which accumulates in tissues as crocodiles lie under water, acts as an allosteric regulator in these animals.[162,163] It allows the animals to more completely utilize the $O_2$ from the hemoglobin and to remain under water longer. The Bohr effect, which was considered in the preceding section, can be viewed as resulting from the action of **protons** as allosteric effectors that bind to the amino and imidazole groups of the salt linkages. **Carbon dioxide** also acts as a physiological effector in mammalian blood by combining reversibly with $NH_2$-terminal groups of the α and β subunits to form **carbamino** ($-NH-COO^-$) groups (Eq. 7-47).[119,164] It is the N-terminal amino groups rather than lysyl side chain groups that undergo this reaction. Because of their relatively low p$K_a$ values there is a significant




(7-47)

fraction of unprotonated $-NH_2$ groups at the pH of blood. The affinity for $CO_2$ is highest in deoxygenated hemoglobin. Consequently, unloading of $O_2$ is facilitated in the $CO_2$-rich respiring tissues. Hemoglobin carries a significant fraction of $CO_2$ to the lungs, and there the oxygenation of hemoglobin facilitates the dissociation of $CO_2$ from the carbamino groups. Hemoglobin is also one of the major pH buffers of blood.

***Carbon monoxide, cyanide, and nitric oxide.*** A danger to hemoglobin and other heme proteins is posed by competing ligands such as CO, $CN^-$, and NO. All of these are present within organisms and both CO and NO act as hormones. Hemoglobin and myoglobin are partially protected from carbon monoxide by the design of the binding site for $O_2$. The distal imidazole of histidine E7 hydrogen bonds to $O_2$ but not to the nonpolar CO. The site also accommodates the geometry of the bound $O_2$ better than that of CO.[165] Bound CO can be released from hemes by the action of light. Using X-ray diffraction[166,166a] and X-ray absorption measurements[167] at cryogenic temperatures, it has been possible to observe the motions of both the released CO and the heme in myoglobin, motions which may shed light on the normal oxygen transport cycle. Cooperativity in the binding of CO to hemoglobins has been studied extensively,[158,168,169] as has binding to model heme compounds.[170] Cyanide ions bind most tightly and also cooperatively[171–173] to the oxidized $Fe^{3+}$ form, which is called **methemoglobin**.

Nitric oxide is a reactive, paramagnetic gaseous free radical which is formed in the human body and in other organisms by an enzymatic oxidation of L-arginine (Eq. 18-65). Since about 1980, NO has been recognized as a hormone with a broad range of effects

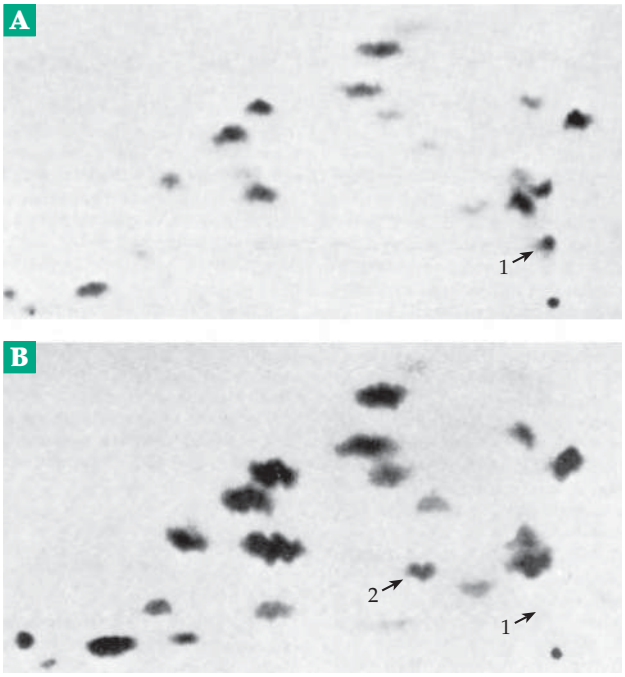



**Figure 7-27** "Fingerprints" of human hemoglobins. The denatured hemoglobin was digested with trypsin and the 28 resulting peptides were separated on a sheet of paper by electrophoresis in one direction (horizontal in the figures; anode to the left) and by chromatography in the other direction (vertical in the figure). The peptides were visualized by spraying with ninhydrin or with specific reagents for histidine or tyrosine residues. Since trypsin cuts only next to lysine, which occurs infrequently, the petide pattern provides a fingerprint, characteristic for any pure protein. (A) The fingerprint of normal adult hemoglobin A. (B) Fingerprint of hemoglobin S (sickle cell hemoglobin). One histidine-containing peptide (1) is missing and a new one (2) is present. This altered peptide contains the first eight residues of the N-terminal chain of the subunit of the protein. From H. Lehmann and R. G. Huntsman, *Man's Haemoglobin*.[187]

(Chapters 11, 18). It binds to the iron of heme groups in either the $Fe^{2+}$ or $Fe^{3+}$ form and also reacts with thiol groups of proteins and small molecules to form *S*-nitrosothiols (R–S–N=O).[174–176] It reacts with the heme iron of myoglobin and hemoglobin and, by transfer of one electron, can oxidize the iron of hemoglobin to the $Fe^{3+}$ methemoglobin with formation of the nitroxyl ion $NO^-$.[177,178] This reaction may be a major cause of methemoglobin formation.

One of the major effects of NO is to induce the relaxation of smooth muscle of blood vessels, an important factor in the regulation of blood pressure. Hemoglobin can carry NO both on its heme and on the thiol group of cysteine β93.[174,175] The affinity for NO is high in the T state and low in the R state. This allows hemoglobin to carry NO from the lungs to tissues, where it can be released and participate in the regulation of blood pressure.[174,179]

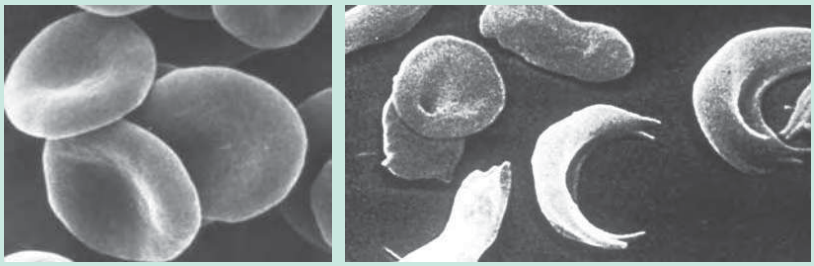
A cytoplasmic hemoglobin of the clam *Lucina pectinata* has evolved to carry oxygen to symbiotic chemoautotrophic bacteria located within cells of the host's gills. It is also readily oxidized to the $Fe^{3+}$ methemoglobin form which binds **sulfide ions** extremely tightly[180–182] and is thought to transport sulfide to the bacteria.

## 2. Abnormal Human Hemoglobins

Many alterations in the structure of hemoglobin have arisen by mutations in the human population. It is estimated that one person in 20 carries a mutation that will cause a hemoglobin disorder in a homozygote.[183] There are also many unrecognized and harmless substitutions of one amino acid for another. However, substitutions near the heme group often adversely affect the binding of oxygen and substitutions in the interfaces between subunits may decrease the cooperative interaction between subunits.[184] One of the most common and serious abnormal hemoglobins is **hemoglobin S**, which is present in individuals suffering from **sickle cell disease** (see Box 7-B). In Hb S, glutamic acid 6 of the β chain is replaced by valine. Replacement of the same amino acid by lysine leads to **Hb C**[185] and is associated with a mild disease condition . A few of the many other substitutions that have been studied are indicated in Fig. 7-23. The locations of the defects in the hemoglobin structure have been established with the aid of protein "fingerprinting" (Fig. 7-27).

A group of serious defects are represented by the **hemoglobins M**. Only heterozygotic individuals survive. Their blood is dark because in Hb M the iron in half of the subunits is oxidized irreversibly to the ferric state. The resulting methemoglobin is present in normal blood to the extent of about 1%. While normal methemoglobin is reduced by a **methemoglobin reductase** system (Box 15-H), methemoglobins M cannot be reduced. All of the five hemoglobins M result from substitutions near the heme group. In four of them, one of the heme-linked histidines (either F-8 or E-7) of either the α or the β subunits is substituted by tyrosine. In the fifth, valine 67 of the β chains is substituted by glutamate. The two hemoglobins M that carry substitutions in the α subunits ($M_{Boston}$ and $M_{Iwate}$) are frozen in the T (deoxy) conformation and therefore have low oxygen affinities and lack cooperativity.

## BOX 7-B SICKLE CELL DISEASE, MALARIA, AND BLOOD SUBSTITUTES



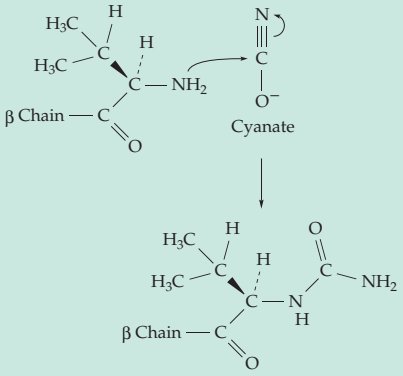
Left: Normal erythrocytes, © Biophoto Associates, Photo Researchers.
Right: Sickled erythrocytes, © Nigel Calder.

Many persons, especially if they are of west African descent, suffer from the crippling and often lethal sickle cell disease.[a,b] In 1949, Pauling, Itano, and associates discovered that hemoglobin from such individuals migrated unusually rapidly upon electrophoresis.[c] Later, Ingram devised the method of protein fingerprinting illustrated in Fig. 7-27 and applied it to hemoglobin.[d] He split the hemoglobin molecule into 15 tryptic peptides which he separated by electrophoresis and chromatography. From these experiments the abnormality in sickle cell hemoglobin (hemoglobin S; Hb S) was located at position 6 in the β chain (see Fig. 7-23). The glutamic acid present in this position in hemoglobin A was replaced by valine in Hb S. This was the first instance in which a genetic disease was traced directly to the presence of a single amino acid substitution in a specific protein. The DNA of the normal gene for the β globin chain has since been sequenced and found to have the glutamic acid codon GAG at position 6. A single base change to GTG (see Table 5-5) causes the sickle cell mutation. Persons homozygous for this altered gene have sickle cell disease, while the much more numerous heterozygotes have, at most, minor problems.

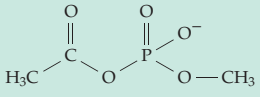
When HbS is deoxygenated it tends to "crystallize" in red blood cells, which contain 33% by weight hemoglobin. The crystallization (actually gel formation) distorts the cells into a sickle shape and these distorted corpuscles are easily destroyed, leading to anemia. The introduction of the hydrophobic valine residue in Hb S at position 6 near the end of the molecule helps form a new bonding domain by which the hemoglobin tetramers associate to form long semicrystalline microfilamentous arrays.[b,e–g]

Why is there such a high incidence of the sickle cell gene, estimated to be present in three million Americans? The occurrence and spread of the gene in Africa was apparently the result of a balance between its harmful effects and a beneficial effect under circumstances existing there. The malaria parasite, the greatest killer of all time, lives in red blood cells during part of its life cycle (see Fig. 1-9). Red cells that contain Hb S as well as Hb A are apparently less suitable than cells containing only Hb A for growth of the malaria organism. Thus, heterozygotic carriers of the sickle cell gene survived epidemics of malaria but at the price of seeing one-fourth of their offspring die of sickle cell disease.

What is the outlook for the many (50,000 in the United States alone) sufferers of sickle cell disease today? Careful medical care, including blood transfusion, can prolong life greatly[h] and intense efforts are under way to find drugs that will prevent Hb S from crystallizing.[i] The problem arises from a hydrophobic interaction of valine B6 with phenylalanine B85 and leucine B88 of another molecule in the filaments of Hb S. The latter two residues are on the outside surface of helix F (see Fig. 7-23). It is difficult to modify one of these residues chemically but various alterations at the nearby N-termini of the β chains do inhibit sickling. Cyanate does so by specifically carbamoylating these amino groups. However, although it was tested in humans, cyanate is too toxic for use.[j] Another approach employs an aldehyde that will



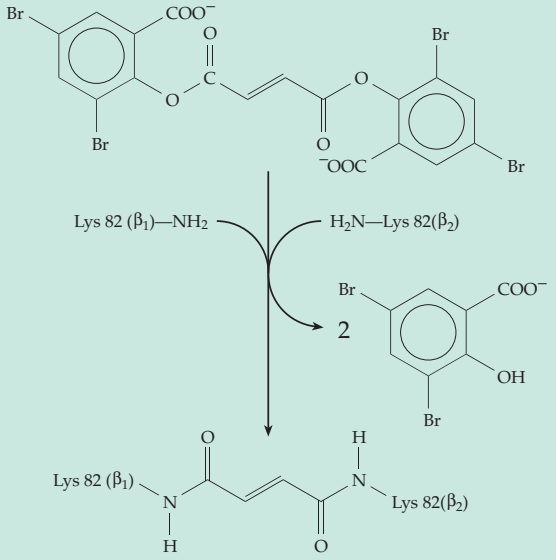

form Schiff bases (Eq. 13-4)[k] with the same amino groups.[k,l] A third approach is to use an acylating reagent. For example, methylacetyl phosphate[m] acetylates the same β Lys 82 amino groups that react with bisphosphoglycerate and with cyanate.



Aspirin (2-acetoxybenzoic acid) is also a mild acetylating reagent and "two-headed" aspirins such

**BOX 7-B (continued)**

as the following react specifically to crosslink the hemoglobin β chains.




These compounds bind into the bisphosphoglycerate binding site (Fig. 7-26) and prevent the chains from spreading apart as far as they normally do in the deoxy (T) state. Since it is only the latter that crystallizes, the compounds have a powerful antisickling action.[n,o] Various other crosslinking reagents have been developed and more than one could be used together.[p–r] New drugs that serve as allosteric modifiers in the same fashion as bisphosphoglycerate may also be useful.[s]

A fouth approach to treatment of sickle cell disease is gene therapy. This might allow patients to produce, in addition to Hb S, an engineered hemoglobin with compensating mutations that would mix with the Hb S and prevent gelling.[t,u,v] This is impractical at present but there is another approach. Persons with sickle cell disease sometimes also have the disorder of hereditary persistence of fetal hemoglobin. They continue to make Hb F into adulthood. Great amelioration of sickle cell disease is observed in patients with 20–25% Hb F.[t] Hydroxyurea stimulates a greater production of Hb F and in patients with hereditary persistence of Hb F hydroxyurea may raise its level in erythocytes to ~50% of the total hemoglobin.[w]

Crosslinking of the *alpha* chains of normal deoxyhemoglobin through lysines 99 yields a hemoglobin with normal oxygen-binding behavior and an increased stability.[o] It makes a practical emergency blood substitute, whereas unmodified hemoglobin is unsatisfactory.[v] Unless encapsulated in an erythrocyte the hemoglobin tetramers tend to dissociate to dimers, losing cooperativity and escaping through kidneys. Suitable crosslinking helps to solve this problem.[x,y,z] Both α and β chains can be produced from cloned genes and reassembled to form hemoglobin.[aa,bb] This will probably allow genetic engineering to form more stable but suitably cooperative hemoglobins that can be used to avoid hazards of transmission of viruses by transfusion.


[a] Weatherall, D. J., Clegg, J. B., Higgs, D. R., and Wood, W. G. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3417–3484, McGraw-Hill, New York

[b] Harrington, D. J., Adachi, K., and Royer, W. E., Jr. (1998) *J. Biol. Chem.* **273**, 32690–32696

[c] Strasser, B. J. (1999) *Science* **286**, 1488–1490

[d] Ingram, V. M. (1957) *Nature (London)* **180**, 326–328

[e] Cretegny, I., and Edelstein, S. J. (1993) *J. Mol. Biol.* **230**, 733–738

[f] Padlan, E. A., and Love, W. E. (1985) *J. Biol. Chem.* **260**, 8272–8279

[g] Cao, Z., and Ferrone, F. A. (1996) *J. Mol. Biol.* **256**, 219–222

[h] Acquaye, C., Wilchek, M., and Gorecki, M. (1981) *Trends Biochem. Sci.* **6**, 146–149

[i] Klotz, I. M., Haney, D. N., and King, L. C. (1981) *Science* **213**, 724–731

[j] Harkness, D. R. (1976) *Trends Biochem. Sci.* **1**, 73–

[k] Acharya, A. S., Sussman, L. G., and Manning, J. M. (1983) *J. Biol. Chem.* **258**, 2296–2302

[l] San George, R. C., and Hoberman, H. D. (1986) *J. Biol. Chem.* **261**, 6811–6821

[m] Ueno, H., Pospischil, M. A., and Manning, J. M. (1989) *J. Biol. Chem.* **264**, 12344–12351

[n] Walder, J. A., Walder, R. Y., and Arnone, A. (1980) *J. Mol. Biol.* **141**, 195–216

[o] Chatterjee, R., Welty, E. V., Walder, R. Y., Pruitt, S. L., Rogers, P. H., Arnone, A., and Walder, J. A. (1986) *J. Biol. Chem.* **261**, 9929–9937

[p] Benesch, R. E., and Kwong, S. (1988) *Biochem. Biophys. Res. Commun.* **156**, 9–14

[q] Kluger, R., Wodzinska, J., Jones, R. T., Head, C., Fujita, T. S., and Shih, D. T. (1992) *Biochemistry* **31**, 7551–7559

[r] Jones, R. T., Shih, D. T., Fujita, T. S., Song, Y., Xiao, H., Head, C., and Kluger, R. (1996) *J. Biol. Chem.* **271**, 675–680

[s] Abraham, D. J., Wireko, F. C., Randad, R. S., Poyart, C., Kister, J., Bohn, B., Liard, J.-F., and Kunert, M. P. (1992) *Biochemistry* **31**, 9141–9149

[t] McCune, S. L., Reilly, M. P., Chomo, M. J., Asakura, T., and Townes, T. M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9852–9856

[u] Cole-Strauss, A., Yoon, K., Xiang, Y., Byrne, B. C., Rice, M. C., Gryn, J., Holloman, W. K., and Kmiec, E. B. (1996) *Science* **273**, 1386–1388

[v] May, C., Rivella, S., Callegari, J., Heller, G., Gaensler, K. M. L., Luzzatto, L., and Sadelain, M. (2000) *Nature (London)* **406**, 82–86

[w] Eaton, W. A., and Hofrichter, J. (1995) *Science* **268**, 1142–1143

[x] Snyder, S. R., Welty, E. V., Walder, R. Y., Williams, L. A., and Walder, J. A. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7280–7284

[y] Dick, L. A., Heibel, G., Moore, E. G., and Spiro, T. G. (1999) *Biochemistry* **38**, 6406–6410

[z] Manjula, B. N., Malavalli, A., Smith, P. K., Chan, N.-L., Arnone, A., Friedman, J. M., and Acharya, A. S. (2000) *J. Biol. Chem.* **275**, 5527–5534

[aa] Yamaguchi, T., Pang, J., Reddy, K. S., Witkowska, H. E., Surrey, S., and Adachi, K. (1996) *J. Biol. Chem.* **271**, 26677–26683

[bb] Jeong, S. T., Ho, N. T., Hendrich, M. P., and Ho, C. (1999) *Biochemistry* **38**, 13433–13442

In hemoglobins Rainier and Nancy the usually invariant C-terminal tyrosine 145 of the β chains is substituted by cysteine and by aspartate, respectively. Oxygen affinity is high and cooperativity is lacking.[186] Hemoglobin Kansas, in which the $\beta^{102}$ asparagine is substituted by threonine, also lacks cooperativity and has a very low oxygen affinity, while hemoglobin Richmond, in which the same amino acid is substituted by lysine, functions normally. In hemoglobin Creteil the β89 serine is replaced by asparagine with the result that the adjacent C-terminal peptide carrying tyrosine 145 becomes disordered. In hemoglobin Hiroshima the C-terminal histidine in the β chain is replaced by aspartic acid. This histidine is one that donates a Bohr proton and in the mutant hemoglobin the oxygen affinity is increased 3-fold and the Bohr effect is halved.[188] In hemoglobin Suresnes the C-terminal arginines 141 of the α chains are replaced by histidine with loss of one of the anion-binding sites mentioned in Section 3.[189]

### 3. Comparative Biochemistry of Hemoglobin

Even within human beings there are several hemoglobins. In addition to myoglobin, a brain protein neuroglobin,[189a] and adult hemoglobin A (**Hb A**, $\alpha_2\beta_2$), there is a minor hemoglobin $A_2$ ($\alpha_2\delta_2$). Prior to birth the blood contains **fetal hemoglobin**, also called hemoglobin F (**Hb F**, $\alpha_2\gamma_2$). In the presence of 2,3-bisphosphoglycerate Hb F has a 6-fold higher oxygen affinity than Hb A as befits its role in obtaining oxygen from the mother's blood.[190–192] Hemoglobin F disappears a few months after birth and is replaced by Hb A. Each of the hemoglobins differs from the others in amino acid sequence.

In other species the amino acid composition of hemoglobins varies more, as do the interactions between subunits. Hemoglobins and myoglobins are found throughout the animal kingdom and even in plants.[192a] The **leghemoglobins**[193–195] are formed in root nodules of legumes and are involved in nitrogen fixation by symbiotic bacteria. Other hemoglobins apparently function in the roots of plants.[196–197] Hemoglobins or myoglobins are found in some cyanobacteria[198] and in many other bacteria.[197,199] The globin fold of the polypeptide is recognizable in all of these.[200]

The quaternary structure of hemoglobin also varies. Myoglobin is a monomer, as is the leghemoglobin. Hemoglobin of the sea lamprey dissociates to monomers upon oxygenation.[201,201a] The clam *Scapharca inaequivalvis* has a dimeric hemoglobin that binds $O_2$ cooperatively even though the interactions between subunits are very different from those in mammalian hemoglobins.[202–204] Hemoglobin of the nematode *Ascaris* is an octamer.[205,206] It has puzzling properties, including a very high affinity for $O_2$ and a slow dissociation rate. The distal His E-7 is replaced by glutamine, which has a hydrogen-bonding ability closely similar to that of histidine.

Earthworms,[207] polychaete worms,[208,209] and leeches[210] have enormous hemoglobin molecules consisting of as many as 144 globin chains arranged into 12 dodecamers and held together by 36–42 linker chains. These hemoglobins are often called **erythrocruorins**. In a few families of polychaetes chloroheme (Fig. 16-5) substitutes for heme and the proteins are called **chlorocruorins**.[208]

What is common to all of the hemoglobins? The same folding pattern of the peptide chain is always present. The protein is always wrapped around the heme group in an identical or very similar manner. In spite of this striking conservation of overall structure, when animal hemoglobins are compared, *there are only ten residues that are highly conserved*. They are indicated in Fig. 7-23 by the boxes. The two glycines (or alanine) at B-6 and E-8 are conserved because the close contact between the B and E helices does not permit a larger side chain. Proline C-2 helps the molecule turn a corner. Four of the other conserved residues are directly associated with the heme group. Histidine E-7 and His F-8 are the "heme-linked" histidines. Tyrosine HC-2, as previously mentioned, plays a role in the cooperativity of oxygen binding. Only Lys H-9 is on the outside of the molecule. The reasons for its conservation are unclear.[211] When sequences from a broader range of organisms were determined five residues (see Fig. 7-23) were found to be highly conserved; *only two are completely conserved.* These are His F-8 and Phe CD-1, which binds the heme noncovalently.[212] Hemoglobins are not the only biological oxygen carriers. The two-iron **hemerythrins** (Fig. 16-20) are used by a few phyla of marine invertebrates, while the copper containing **hemocyanins** (Chapter 16, Section D,4) are used by many molluscs and arthropods.

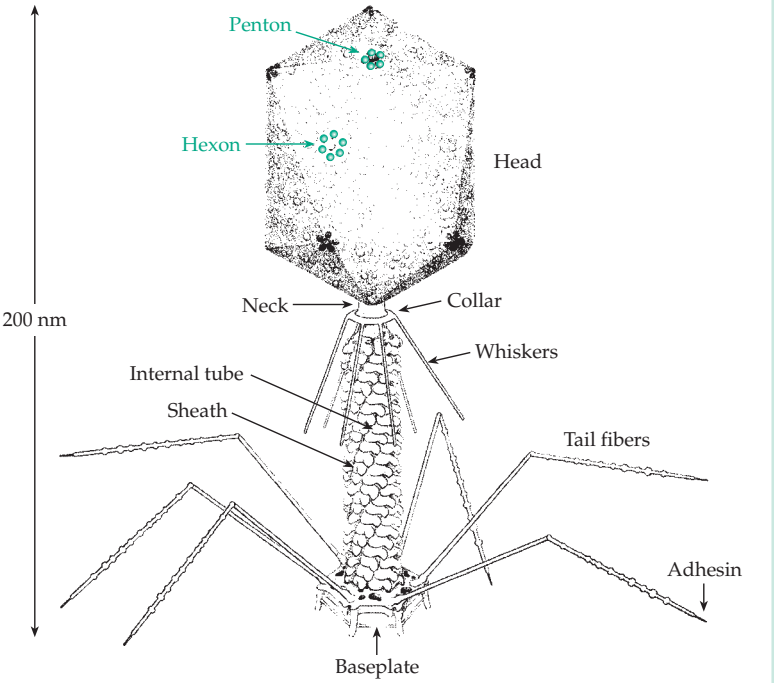
## E. Self-Assembly of Macromolecular Structures

While it is easy to visualize the assembly of oligomeric proteins, it is not as easy to imagine how complex objects such as eukaryotic cilia (Fig. 1-8) or the sarcomeres of muscle (Fig. 19-6) are formed. However, study of the assembly of bacteriophage particles and other small biological objects has led to the concepts of **self-assembly** and **assembly pathways**, concepts that are now applied to every aspect of the architecture of cells.

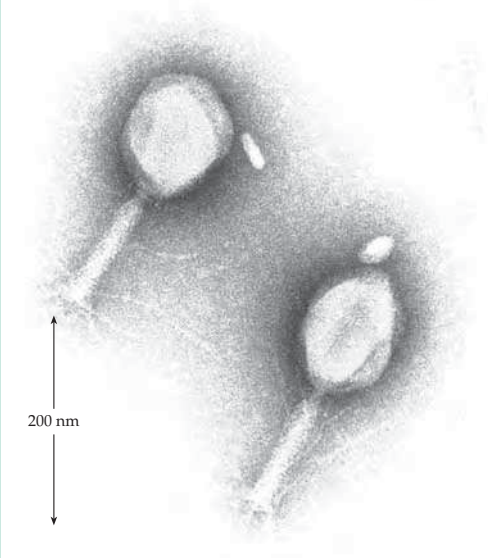
### 1. Bacteriophages

A remarkable example of self-assembly is that of the T-even phage (Box 7-C).[213–215] From genetic analysis (Chapter 26) at least 22 genes are known to be required for formation of the heads, 21 genes for the tails, and 7 genes for the tail fibers. Many of these genes encode

## BOX 7-C THE T-EVEN BACTERIOPHAGES




Drawing courtesy of F. Eiserling and the American Society for Microbiology.[a]




Bacteriophage T4. ×240,000
Micrograph courtesy of Tom Moninger.

Among the most remarkable objects made visible by the electron microscope are the T-even bacteriophage (T2, T4, and T6) which attack *E. coli*.[a–e] While it is not often evident how a virus gains access to a cell, these "molecular syringes" literally inject their DNA through a hole dissolved in the cell wall of the host bacterium. The viruses, of length ~200 nm and mass ~225 x $10^6$ Da each, contain 130 x $10^6$ Da of DNA in a 100 x 70 nm head of elongated icosahedral shape. The head surface appears to be formed from ~840 copies of a 45-kDa protein known as gp23 (gene product 23; it is encoded by gene 23; see map in Fig. 26-2). These protein molecules are arranged as 140 hexamers (hexons) and together with ~55 copies of protein gp24, arranged as 11 pentamers (pentons), make up the bulk of the shell.[f] The head also contains at least nine other proteins, including three internal, basic proteins that enter the bacterium along with the DNA. Additional proteins form the **neck**, **collar**, and **whiskers**. The phage **tail**, which fastens to the collar via a **connector** protein,[g] contains an **internal tube** with a 2.5-nm hole, barely wide enough to accommodate the flow of the DNA molecule into the bacterium. The tube is made up of 144 subunits of gp19. The 8 x $10^6$-Da **sheath** that surrounds the tail tube is made up of 144 subunits of gp18, each of mass 55 kDa, arranged in the form of 24 rings of six subunits each.[h,i] The sheath has contractile properties. After the virus has become properly attached to the host it shortens from ~80 to ~30 nm, forcing the inner tube through a hole etched in the wall of the bacterium. At the end of the tail is a **baseplate**, a hexagonal structure bearing six short **pins**, each a trimer of a 55 kDa zinc metalloprotein.[j] One of the ten proteins known to be present in the baseplate is the enzyme T4 lysozyme (Chapter 12). The baseplate also contains six molecules of the coenzyme 7,8-dihydropteroyl-hexaglutamate (Chapter 15, Section D).

Six elongated, "jointed" **tail fibers** are attached to the baseplate. The proximal segment of each fiber is a trimer of the 1140-kDa protein gp34. A globular domain attaches it to the baseplate. The distal segment is composed of three subunits of the 109-kDa gp37, three subunits of the 23-kDa gp36, and a single copy of the 30-kDa gp35.[k] The C-terminal ~140 residues of the 1026-residue gp37 are the specific **adhesin** that binds to a lipopolysaccharide of *E. coli* type B cells or to the outer membrane protein OmpC (Chapter 8).[l] Among the smaller molecules present in the virus are the polyamines **putrescine** and **spermidine** (Chapter 24), which neutralize about 30% of the basic groups of the DNA.

How is infection by a T-even virus initiated? Binding of the tail fibers to specific receptor sites on the bacterial surface triggers a sequence of conformational changes in the fibers, baseplate, and

## BOX 7-C THE T-EVEN BACTERIOPHAGES (continued)

sheath. The lysozyme is released from the baseplate and etches a hole in the bacterial cell wall. Contraction of the sheath is initiated at the baseplate and continues to the upper end of the sheath. The tail tube is forced into the bacterium and the DNA rapidly flows through the narrow hole into the host cell.

During contraction the subunits of the sheath undergo a remarkable rearrangement into a structure containing 12 larger rings of 12 subunits each.[h] Thus, a kind of mutual "intercalation" of subunits occurs. In its unidirectional and irreversible nature the shortening of the phage tail differs from the contraction of muscle. The protein subunits of the sheath seem to be in an unstable high energy state when the tail sheath of the phage is assembled. The stored energy remains available for later contraction.


[a] Mathews, C. K., Kutter, E. M., Mosig, G., and Berget, P. B., eds. (1983) *Bacteriophage $T_4$*, Am. Soc. Microbiology, Washington, D.C.
[b] Wood, W. B., and Edgar, R. S. (1967) *Sci. Am.* **217**(Jul), 60–74
[c] Mathews, C. K. (1971) *Bacteriophage Biochemistry*, Van Nostrand-Reinhold, Princeton, New Jersey
[d] Voyles, B. A. (1993) *The Biology of Viruses*, Mosby, St. Louis, Missouri
[e] Tikhonenko, A. S. (1970) *Ultrastructure of Bacterial Viruses*, Plenum, New York
[f] Branton, D., and Klug, A. (1975) *J. Mol. Biol.* **92**, 559–565
[g] Cerritelli, M. E., and Studier, F. W. (1996) *J. Mol. Biol.* **258**, 299–307
[h] Moody, M. F., and Makowski, L. (1981) *J. Mol. Biol.* **150**, 217–244
[i] Müller, D. J., Engel, A., Carrascosa, J. L., and Vélez, M. (1997) *EMBO J.* **16**, 2547–2553
[j] Zorzopulos, J., and Kozloff, L. M. (1978) *J. Biol. Chem.* **253**, 5543–5547
[k] Cerritelli, M. E., Wall, J. S., Simon, M. N., Conway, J. F., and Steven, A. C. (1996) *J. Mol. Biol.* **260**, 767–780
[l] Tétart, F., Repoila, F., Monod, C., and Krisch, H. M. (1996) *J. Mol. Biol.* **258**, 726 – 731


sequences of proteins that are incorporated into the mature virus, but several specify enzymes needed in the assembly process. Several mutant strains of the viruses are able to promote synthesis of all but one of the structural proteins of the virion. Proteins accumulating within these defective bacterial hosts have no tendency to aggregate spontaneously. However, when the missing protein (synthesized by bacteria infected with another strain of virus) is added complete virus particles are formed rapidly. Investigations resulting from this and other observations have led to the conclusion that during assembly *each different protein is added to the growing aggregate in a strictly specified sequence or assembly pathway.* The addition of each protein creates a binding site for the next protein. In some cases the protein that binds is an enzyme that cuts off a piece from the growing assembly of subunits and thereby creates a site for the next protein to bind.

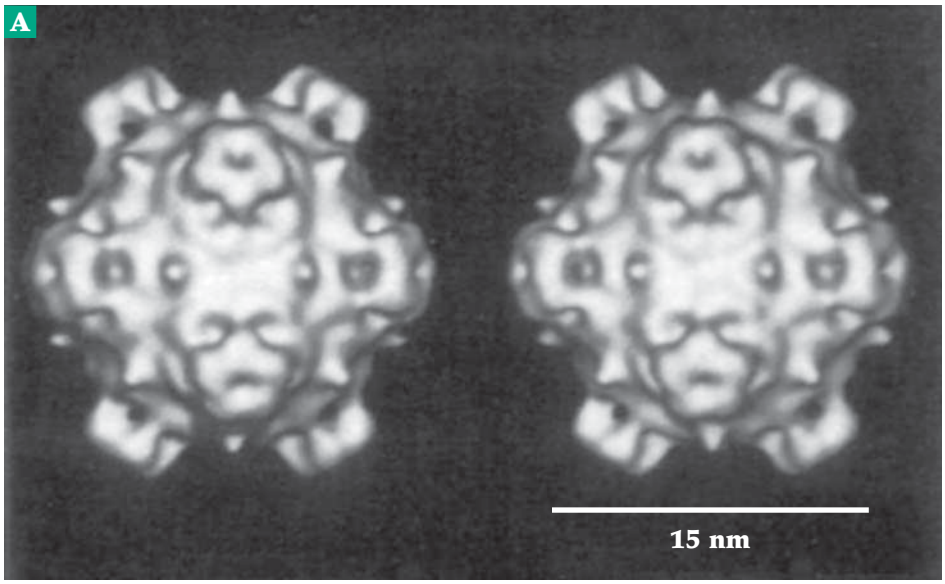
Before considering this complex process further, let's look at the assembly of simpler filamentous bacteriophages (Fig. 7-7) and bacterial pili (Fig. 7-9). The filamentous bacteriophages are put together from hydrophobic protein subunits and DNA. After their synthesis the protein subunits are stored within the cytoplasmic membrane of the infected bacteria.[216,217] These small, largely α-helical rods can easily fit within the membrane and remain there until a DNA molecule also enters the membrane.[218,219] Two additional proteins (gene I proteins) also enter the membrane. One has a 348-residue length, while the second is a 107-residue protein formed by translational initiation at a later point in the DNA sequence of the gene. These proteins help to create an assembly site at a place where the inner and outer membrane of the host bacterium are close together.[220] It isn't clear how the process is initiated, but it is likely that each subunit of the viral coat contains a nucleotide binding site that interacts with the DNA. Adjacent sides of the subunits are hydrophobic and interact with other subunits to spontaneously coat the DNA. As the rod is assembled, the hydrophobic groups become "buried." It is postulated that the remaining side chain groups on the outer surface of the virus are hydrophilic and that the formation of this hydrophilic rod provides a driving force for automatic extrusion of the phage from the membrane.[216]

Bacterial pili appear to be extruded in a similar manner. They arise rapidly and may possibly be retracted again into the bacterial membrane. The P pilus in Fig. 7-9A is made up of subunits PapA, G, F, E, and K which must be assembled in the correct sequence. A chaperonin PapD is also required as is an "**usher protein**," PapC,[50] and also the disulfide exchange protein DsbA (Chapter 10). DsbA helps PapD to form the correct disulfide bridges as it folds and PapD binds and protects the various pilus subunits as they accumulate in the periplasmic space of the host. The usher protein displaces the chaperonin PapD and "escorts" the subunits into the membrane where the extrusion occurs.[50,55]

Because eicosahedra are regular geometric solids and the faces can be made up of hexons and pentons of identical subunits, it might seem that self-assembly of eicosahedral viruses would occur easily. However, the subunits usually must be able to assume three or more different conformations and the shells can easily be assembled incorrectly. Several stategies are em-

ployed to avert this problem.[221,222] Some viruses assemble an empty shell into which the DNA flows, but many others first form an internal **scaffolding** or **assembly core** around which the shell is assembled. An external scaffold may also be needed.[223] The RNA virus MS2 forms its T=3 capsid by using the RNA molecule as the assembly core.[224] Other viruses may have one or more core proteins which dissociate from the completed shell or are removed by the action of proteases. This is a feature of the small ϕX phage (Fig. 7-28),[225,226,230] the tailed phages,[227,228] and double-stranded RNA viruses including human reoviruses.[229] Bacteriophage PRD1, another virus of *E. coli* and *Salmonella typhimurium*, has a membrane inside the capsid apparently playing a role in assembly.[232]

A general concept that seems to hold in all cases is one of "**local rule**." Several conformers of a virus subunit may equilibrate within a cell. However, *they*



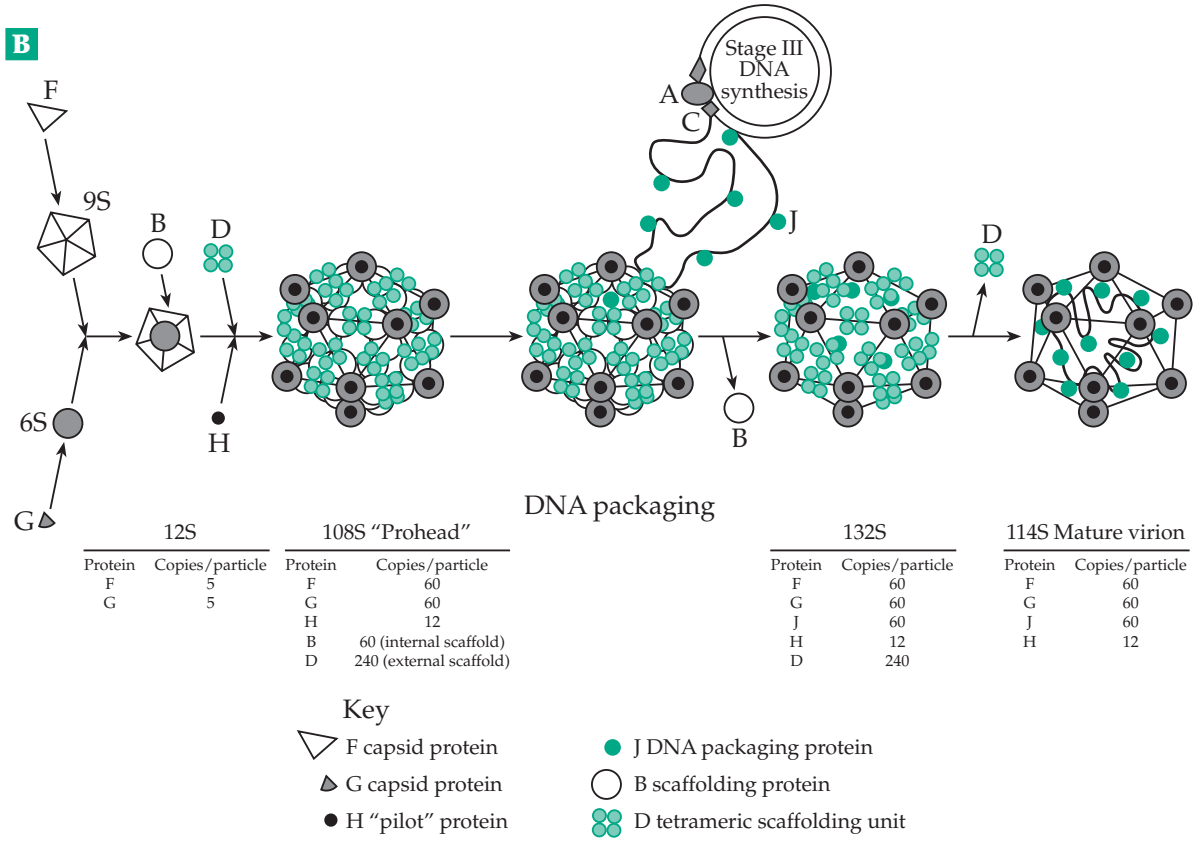




**Figure 7-28** (A) Stereoscopic view of the ϕX174 114 S mature virion viewed down a twofold axis after a cryoelectron microscopy reconstruction. From McKenna *et al*.[225] (B) Morphogenesis of ϕX174 (based on a report of Hayashi *et al*.[231]). Proteins A and C are required for DNA synthesis. Drawing from McKenna *et al*.[225] Courtesy of Michael G. Rossmann.

*can associate only through surfaces that are complementary*. A conformer that allows pentons to form cannot assemble into a hexon and only certain combinations of other conformers can give rise to hexons, etc.[233] If there is only one conformation and the shape is right a T=1 shell will be formed. If there are three conformers a T= 3 shell may arise. Another generalization is that in most cases the **procapsid** or **prohead** that is formed initially is fragile. Subunits may still be undergoing conformational changes. However, a final conformational alteration, which may include chain cleavage by a protease, usually occurs. This often expands the overall dimensions of the capsid and creates new intersubunit interactions which greatly strengthen the mature capsid.[234] Scaffolding proteins are then removed and DNA or RNA enters the capsid, again in a precise sequence. There are many variations and the detail that is known about virus assembly is far too great to describe here.

Figure 7-29 illustrates the assembly pathway for the very small ϕX174, a T=1 virus. The major capsid protein F is a 426-residue eight-stranded β-barrel. The 175-residue G protein forms pentameric spikes while 60 copies of the internal scaffolding protein and 240 copies of the external scaffolding protein D and 12 copies of the pilot protein are required to form the prohead. The single-stranded DNA enters along with 60 copies of a DNA packaging protein J.

Assembly of the tailed bacteriophages (Box 7-C) is even more complex. The genome of the viruses is large. The 166-kb circular dsDNA of phage T4 contains ~250 genes, many of which encode proteins of the virion or enzymes or chaperonins needed in assembly. The assembly pathway for the bacteriophage heads requires at least 22 gene products.[235] Seven of these form the assembly core which serves as a scaffolding around which the 840 copies of gp23 and 55 copies of gp24 (see Box 7-C) are added to give the elongated icosahedral prohead I. Most of the internal proteins are then dissolved by proteases, one of which is the phage-specified gp21. A protease also cuts a piece from each molecule of gp23 to form the smaller gp23*, the major protein of the mature **prohead II**. This cleavage also triggers the conformational change leading to head expansion. The empty proheads are now filled with DNA in a process which is assisted by another series of catalytic proteins.

The T4 phage tail is assembled in a separate sequence. Six copies of each of three different proteins form a "hub" with hexagonal symmetry (Fig. 7-29). In another assembly sequence, seven different proteins form wedge-shaped pieces, six of which are then joined to the hub to form the hexagonal baseplate. Two more proteins then add to the surface of the base plate and activate it for the growth of the internal tail tube. Only after assembly of the internal tube begins does the sheath

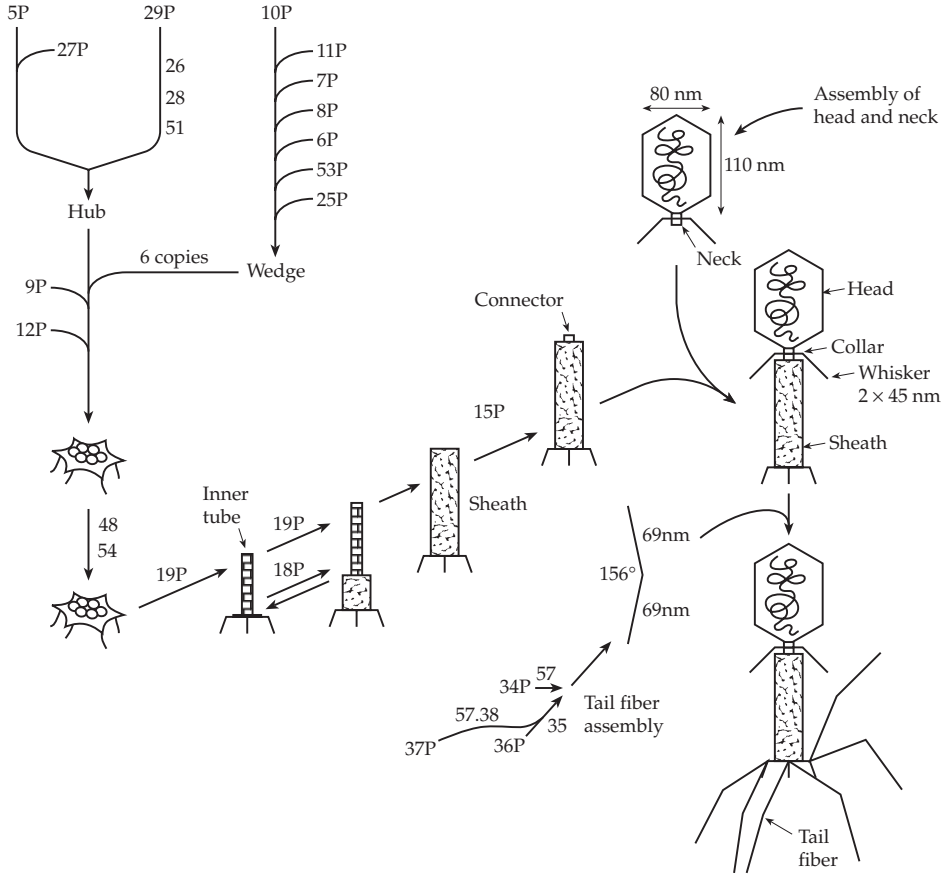



**Figure 7-29** Assembly sequence for bacteriophage T4 with details for the tail. The numbers refer to the genes in the T4 chromosome map (Fig. 26-2). A "P" after the number indicates that the protein gene product is incorporated into the phage tail. Other numbers indicate gene products that are thought to have essential catalytic functions in the assembly process. Adapted from King and Mykolajewycz[236] and Kikuchi and King.[214]

begin to grow, and only when both of these tubular structures have reached the correct length, is a cap protein placed on top. The DNA-filled head is then attached by a special connector and only then do the tail fibers, which have been assembled separately, join at the opposite end.

How can each step in this complex assembly process set the stage for the next step? Apparently the structure of each newly synthesized protein monomer is stable only until a specific interaction with another protein takes place. The binding energy of this interaction is sufficient to induce a conformational alteration that affects a distant part of the protein surface and generates complementarity toward a binding site on the next protein that is to be added. Every one of the baseplate proteins must have such self-activating properties! Sometimes proteolytic cleavage of a subunit is required. If it occurs at an appropriate point in the sequence it provides thermodynamic drive for the assembly process.

The induction of a change in one protein by interaction with another protein is a phenomenon that is met also in the construction of microtubules, ribosomes, cilia, and myofibrillar assemblies of muscle. It is basic to the assembly of the many labile but equally real cascade systems of protein–protein interactions such as that involved in the clotting of blood (Chapter 12) and signaling at membrane surfaces.

## 2. "Kringles" and Other Recognition Domains and Motifs

The assemble of either transient or long lasting complexes of proteins is often dependent upon the presence of conserved structural domains of 30–100 residues. A similar domain may occur in many different proteins and often two or more times within a single protein. The sequences within such a domain are homologous, allowing it to be recognized from protein or gene sequences alone.[237,238] Domains are often named after the protein in which they were first discovered. For example, **EGF-like domains** resemble the 53-residue epidermal growth factor. **SH2** and **SH3** domains are src-homology domains, named after **Src** (c-src), the protein encoded by the *src* protooncogene (Table 11-3).[239] The SH2 and SH3 domains are found near the N terminus of this 60-kDa protein. They are also found in many other proteins. An adapter protein called **Grb2**, important in cell signaling, consists of nothing but one SH2 domain and two SH3 domains (Figs. 11-13, 11-14).[240] The SH2 domains bind to phosphotyrosyl side chains of various proteins, while SH3 domains bind to a polyproline motif. Another phosphotyrosyl binding domain, the plekstrin homology or **PH** domain, is named for the protein in which it was discovered. **Kringle** and **apple** describe the appearances of the folded proteins in those domains. Structural domains often function to hold two proteins together or to help anchor them at a membrane surface by binding to specific protein groups, such as phosphotyrosyl, or calcium ions. Table 7-3 lists a few well-known folding domains and Fig. 7-30 shows three-dimensional structures of two of them.

Recognition domains often function transiently. For example, SH2 domains are often found in proteins that interact with phosphotyrosyl groups of "activated" cell surface receptors. The receptors become activated

**TABLE 7-3**
**A Few Well-Known Structural Domains**

| Name | Length in amino acid residues | Specific ligands |
|---|---|---|
| EGF-like[241–244] | ~45 | $Ca^{2+}$ |
| SH2[238,245–247] | ~100 | Phosphotyrosine |
| Structure[248–253] | | |
| SH3[239,246,254] | ~60 | Polyproline, PXXP |
| Structure[255–257] | | |
| PTB[238,258,259] | | Proline-rich sequence |
| PH (plekstrin homology)[260,261] | | Phosphotyrosine |
| Structure[262,263] | | |
| PDZ[264–266] | 80–100 | C-terminal XS / TXV–$COO^-$ |
| Immunoglobulin repeat (Fig. 2-16)[267] | ~100 | |
| Kringles, blood clotting proteins[268–270] | 80-85 | Calcium binding |
| Apple, Blood clotting Factor X[271] | 90 | Calcium binding |
| WW (Trp–Trp)[272] | ~38 | Proline |
| Serine protease[273] | | |
| P (Trefoil)[274] | ~50 | |
| TPR (Tetratrico peptide repeat)[247,275,276] | | |
| ZBD (Zinc-binding domains) | | |
| Zinc finger (Fig. 5-37)[277] | | |
| Others[277–280] | | |

by conformational alteration resulting from the binding. The src protein is a tyrosine kinase and, when activated, uses ATP to phosphorylate tyrosyl groups of other proteins, and using its SH2 domains it will bind to such groups forming and passing an intracellular message to them.[249]

## F. The Cytoskeleton

The cytoplasm of eukaryotic cells contains a complex network of slender rods and filaments that serve as a kind of internal skeleton. The properties of this **cytoskeleton** affect the shape and mechanical properties of cells. For example, the cytoskeleton is responsible

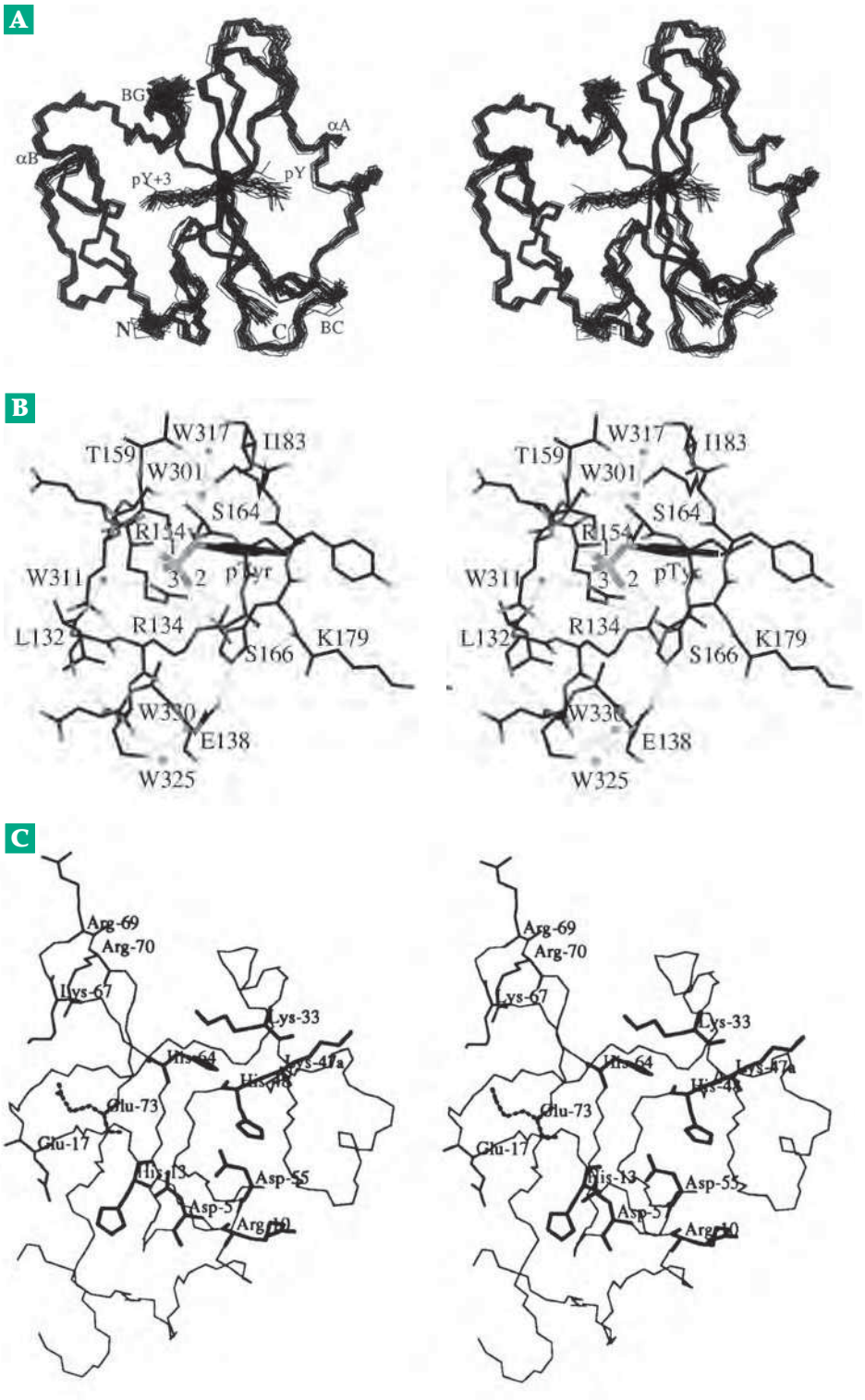



**Figure 7-30** (A) Stereoscopic MolScript view showing 30 superposed solution structures of the SH2 / phosphopeptide complex from protein Shc calculated from NMR data. The N and C termini of the protein as well as the phosphotyrosine (pY) and (pY + 3) residues of the phosphopeptide are indicated. From Zhou *et al*.[252] Courtesy of Stephen W. Fesik. (B) View of the phosphotyrosine side chain of the peptide pYEEI bound to a high-affinity SH2 domain from the human *src* tyrosine kinase called p56$^{lck}$. The phosphate group forms a series of hydrogen bonds with groups in the protein and with water molecules (small dots) and an ion pair with the guanidinium group of R134. From Tong *et al*.[248] Courtesy of Liang Tong. (C) Structure of kringle 2 from human tissue plasminogen activator (see Chapter 12). From de Vos *et al*.[269] Courtesy of Abraham M. de Vos.

for the biconcave disc shape of erythrocytes and for the ameba's ability to rapidly interconvert gel-like and fluid regions of the cytoplasm.[281–283]

Three principal components of the cytoskeleton are **microfilaments** of ~6 nm diameter, **microtubules** of 23–25 nm diameter, and **intermediate filaments** of ~10 nm diameter. A large number of associated proteins provide for interconnections, for assembly, and for disassembly of the cytoskeleton. Other proteins act as **motors** that provide motion. One of these motors is present in **myosin** of muscle. This protein is not only the motor for muscular work but also forms **thick filaments** of 12–16 nm diameter, which are a major structural component of muscle (see Fig. 19-6).

## 1. Intermediate Filaments

In most cells the intermediate filaments provide the scaffolding for the cytoskeleton[284–286] They may account for only 1% of the protein in a cell but provide up to 85% of the protein in the tough outer layers of skin. Intermediate filament proteins are encoded by over 50 human genes[286] which specify proteins of various sizes, structures, and properties. However, all of them have central 300- to 330-residue α-helical regions through which the molecules associate in parallel pairs to form coiled-coil rods with globular domains at the ends (Fig. 7-31). Some of these proteins, such as the **keratin** of skin, are insoluble. Others, including the nuclear **lamins** (Chapter 27)[287] and **vimentin**,[288–289a] dissociate and reform filaments reversibly.

Vimentin is found in most cells and predominates in fibroblasts and other cells of mesenchymal origin. **Desmin** (55-kDa monomer) is found in both smooth and skeletal muscle.[289b,290,290a] In the latter, it apparently ties the contractile myofibrils to the rest of the cytoskeletal network and the individual myofibrils to each other at Z disc (see Fig. 19-6). The **glial filaments** from the astroglial cells of the brain are composed mainly of a single type of 55-kDa subunits of the **glial fibrillar acidic protein** but the **neurofilaments** of mammalian neurons are composed of three distinct subunits of 68-, 150-, and 200-kDa mass.[291–293] The larger subunits have C-terminal tails that are not required for filament formation but which can be phosphorylated and form bridges to neighboring neurofilaments and other cytoskeletal components and organelles. Keratin filaments, which eventually nearly fill the highly differentiated epidermal cells, are also made up of several different subunits.[294] Extensions of the keratin chains are rich in cysteine side chains which form disulfide crosslinkages to adjacent molecules to provide a network that can be dehydrated to form hair and the tough outer layers of skin.[286] Elastin-associated microfibrils are important constituents of elastic tissues of blood vessels, lungs, and skin.[295]

A common architecture of intermediate filaments is a staggered head-to-tail and side-by-side association of pairs of the coiled-coil dimers into 2- to 3-nm protofilaments and further association of about eight protofilaments to form the 10-nm intermediate filaments.[286,290,296]

## 2. Microfilaments

The most abundant microfilaments are composed of fibrous actin (F-actin; Fig. 7-10). The **thin filaments** of F-actin are also one of the two major components of the contractile fibers of skeletal muscle. There is actually a group of closely related actins encoded by a multigene family. At least four vertebrate actins are specific to various types of muscle, while two (β- and γ-actins) are cytosolic.[298,299] Actins are present in all animal cells and also in fungi and plants as part of the cytoskeleton. The microfilaments can associate to

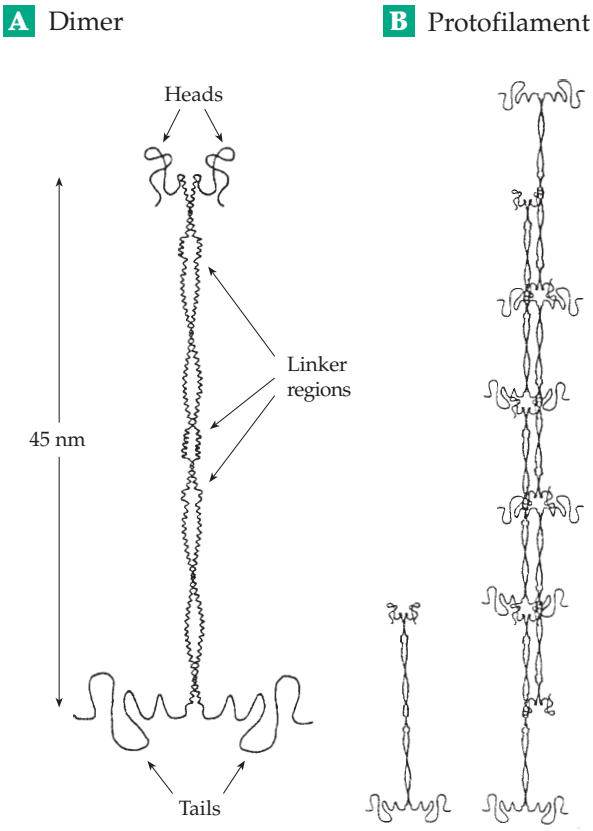



**Figure 7-31** A model for the structure of keratin microfibrils of intermediate filaments. (A) A coiled-coil dimer, 45-nm in length. The helical segments of the rod domains are interrupted by three linker regions. The conformations of the head and tail domains are unknown but are thought to be flexible. (B) Probable organization of a protofilament, involving staggered antiparallel rows of dimers. From Jeffrey A. Cohlberg[297]

form larger arrays and actin often exists as thicker "cables," some of which form the **stress fibers** seen in cultured cells adhering to a glass surface. In the red blood cells the **spectrin–actin** meshwork (Fig. 8-14), which lies directly beneath the plasma membrane, together with the proteins that anchor it to the membrane, form the cytoskeleton.[284,300,301] Its mechanical properties appear to be responsible for the biconcave disc shape of the cell.

The **acrosomal process** of some invertebrate sperm cells is an actin cable that sometimes forms almost instantaneously by polymerization of the actin monomers and shoots out to penetrate the outer layers of the egg during fertilization (Chapter 32). The **stereocilia**, the "hairs" of the hair cells in the inner ear, contain bundles of actin filaments.[302] Motion of the stereocilia caused by sound produces changes in the membrane potential of the cells initiating a nerve impulse. In certain lizards each hair cell contains about 75 stereocilia of lengths up to 30 μm and diameter 0.8 μm and containing more than 3000 actin filaments in a semicrystalline array. Microvilli (Fig. 1-6) contain longitudinal arrays of actin filaments.

In every instance, groups of microfilaments are held together by other proteins. Stress fibers of higher eukaryotes contain the "muscle proteins" **tropomyosin**, **α-actinin**, and myosin (Chapter 19), although the latter is usually not in fibrillar form. **Filamin** (250-kDa) and a 235-kDa protein are associated with actin in platelets.[303] The high-molecular-weight **synemin** crosslinks vimentin and desmin filaments,[304] while the smaller, highly polar **filaggrin** provides a matrix around the keratin filaments in the external layers of the skin.[305]

Postsynthetic modifications of cytoskeletal microfilaments can also occur. For example, epidermal keratin has been found to contain lanthionine, (γ-glutamyllysine) and lysinoalanine, both presumably arising from crosslinkages.[306]

## 3. Microtubules

A prominent component of cytoplasm consists of microtubules which appear under the electron microscope to have a diameter of 24 ± 2 nm and a 13 - to 15-nm hollow core.[307–310] However, the true diameter of a hydrated microtubule is about 30 nm and the microtubule may be further surrounded by a 5–20 nm low density layer of associated proteins. Microtubules are present in the most striking form in the flagella and cilia of eukaryotic cells (Fig. 1-8). The **stable microtubules** of cilia are integral components of the machinery causing their motion (Chapter 19). **Labile microtubules**, which form and then disappear, are often found in cytoplasm in which motion is taking place, for example, in the pseudopodia of the ameba. The mitotic spindle consists largely of microtubules which function in the movement of chromosomes in a dividing cell (Box 7-D and Chapter 26).

Microtubules in the long axons of nerve cells function as "rails" for the "fast transport" of proteins and other materials from the cell body down the axons. In fact, microtubules appear to be present throughout the cytoplasm of virtually all eukaryotic cells (Fig. 7-32) and also in spirochetes.[311] Motion in microtubular systems depends upon motor proteins such as **kinesin**, which moves bound materials toward what is known as the "negative" end of the microtubule,[312] **dyneins** which move toward the positive end.[310] These motor proteins are driven by the Gibbs energy of hydrolysis of ATP or GTP and in this respect, as well as in some structural details (Chapter 19), resemble the muscle protein myosin. Dynein is present in the arms of the microtubules of cilia (Fig. 1-8) whose motion results from the sliding of the microtubules driven by the action of this protein (Chapter 19).

Microtubules are assembled from ~ 55-kDa **tubulins**, which are mixed dimers of α subunits (450 residues) and β subunits (445 residues) with 40% sequence identity. The αβ dimers, whose structure is shown in

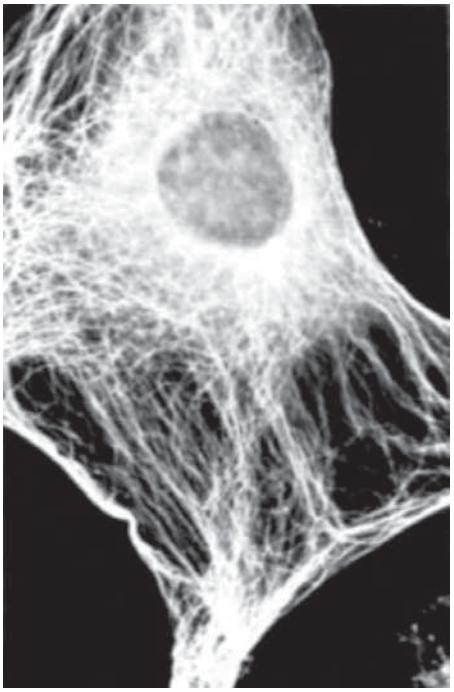
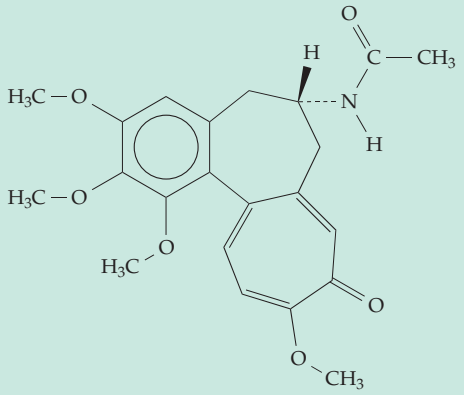


**Figure 7-32** Micrograph of a mouse embryo fibroblast was obtained using indirect immunofluorescence techniques.[313] The cells were fixed with formaldehyde, dehydrated, and treated with antibodies (formed in a rabbit) to microtubule protein. The cells were then treated with fluorescent goat antibodies to rabbit γ-globulins (see Chapter 31) and the photograph was taken by fluorescent light emission. Courtesy of Klaus Weber.

## BOX 7-D MITOSIS, TETRAPLOID PLANTS, AND ANTICANCER DRUGS

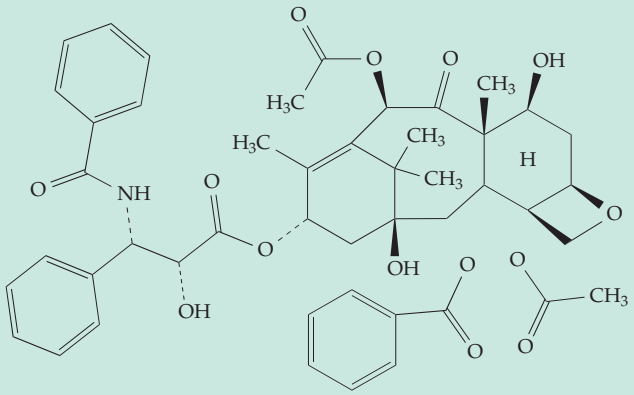
Microtubules in cells undergoing mitosis are the target of several important drugs. One of these is the alkaloid **colchicine** which is produced by various members of the lily family and has been used since ancient Egyptian times for the alleviation of the symptoms of gout.[a,b]



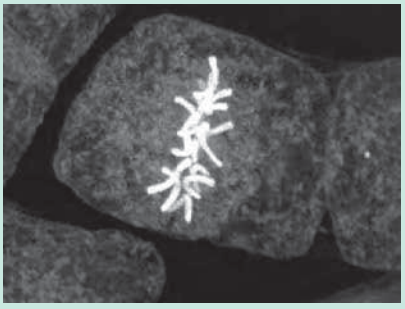
Colchicine

This compound, with its tropolone ring system, binds specifically and tightly and prevents assembly of microtubules, including those of the mitotic spindle.[b] Colchicine forms a complex with soluble tubulin,[c] perhaps a dimeric αβ complex of the two subunits.[d] Dividing cells treated with colchicine appear to be blocked at metaphase (Chapter 26) and daughter cells with a high degree of polyploidy are formed. This has led to the widespread use of colchicine in inducing formation of tetraploid varieties of flowering plants. Similar effects upon microtubules are produced by the antitumor agents **vincristine** and **vinblastine**, alkaloids formed by the common plant *Vinca* (periwinkle), and also by a variety of other drugs.[e]

The more recently discovered Taxol (paclitaxel) was extracted from the bark of the Pacific yew.[f] It *stabilizes* microtubules, inhibiting their disassembly.[g] Taxol also blocks mitosis and causes the cells which fail to complete mitosis to die. Taxol has been synthesized[f] and is a promising drug that is being used in treatment of breast, ovarian, and other cancers.[f] Binding sites for the compound have been located in β tubulin subunits (Fig. 7-33).[h,i] Attempts are being made to develop "taxoids" and other drugs more effective than taxol against cancer cells.[j,k]



Taxol (paclitaxel)



Laser scanning confocal micrograph of chromosomes at metaphase. Courtesy of Tom Moninger

Another group of drugs that bind to microtubules are **benzimidazole** and related compounds. These have been used widely to treat infection by parasitic nematodes in both humans and animals. Unfortunately resistance has developed rapidly. In a nematode that infects sheep a single tyrosine to phenylalanine mutation at position 200 in the β-tubulin subunit confers resistance.[l]


[a] Margulis, T. N. (1974) *J. Am. Chem. Soc.* **96**, 899–902
[b] Chakrabarti, G., Sengupta, S., and Bhattacharyya, B. (1996) *J. Biol. Chem.* **271**, 2897–2901
[c] Panda, D., Daijo, J. E., Jordan, M. A., and Wilson, L. (1995) *Biochemistry* **34**, 9921–9929
[d] Shearwin, K. E., and Timasheff, S. N. (1994) *Biochemistry* **33**, 894–901
[e] Hastie, S. B., Williams, R. C., Jr., Puett, D., and Macdonald, T. L. (1989) *J. Biol. Chem.* **264**, 6682–6688
[f] Nicolaou, K. C., Nantermet, P. G., Ueno, H., Guy, R. K., Couladouros, E. A., and Sorensen, E. J. (1995) *J. Am. Chem. Soc.* **117**, 624–633
[g] Derry, W. B., Wilson, L., and Jordan, M. A. (1995) *Biochemistry* **34**, 2203–2211
[h] Rao, S., He, L., Chakravarty, S., Ojima, I., Orr, G. A., and Horwitz, S. B. (1999) *J. Biol. Chem.* **274**, 37990–37994
[i] Makowski, L. (1995) *Nature (London)* **375**, 361–362
[j] Nicolaou, K. C., Guy, R. K., and Potier, P. (1996) *Sci. Am.* **274**(Jun), 94–98
[k] Kowalski, R. J., Giannakakou, P., and Hamel, E. (1997) *J. Biol. Chem.* **272**, 2534–2541
[l] Kwa, M. S. G., Veenstra, J. G., Van Dijk, M., and Roos, M. H. (1995) *J. Mol. Biol.* **246**, 500–510

Fig. 7-33[314–316] are thought to be packed into an imperfect helix as indicated in Fig. 7-34. The structure can also be regarded as an array of longitudinal protofilaments. Naturally formed microtubules usually have precisely 13 protofilaments and a discontinuity in the helical stacking of subunits as shown in Fig. 7-34. When grown in a laboratory the microtubules usually have 14 protofilaments[317] and rarely 10 or 16 protofilaments with regular helical packing.[318] Microtubules of some moths and also of male germ cells of *Drosophila* have 16-protofilament microtubules without a discontinuity, an architecture that is specified by the geometry of a specific β-tubulin isoform.[319]

Each tubulin dimer binds one molecule of GTP strongly in the α subunit and a second molecule of GTP or GDP more loosely in the β subunit. In this respect, tubulin resembles actin, whose subunits are about the same size. However, there is little sequence similarity. Labile microtubules of cytoplasm can be formed or disassembled very rapidly. GTP is essential for the fast growth of these microtubules and is hydrolyzed to GDP in the process.[320] However, nonhydrolyzable analogs of GTP, such as the one containing the linkage P–$CH_2$–P between the terminal and central phosphorus atoms of the GTP, also support polymerization.[321] Since microtubules have a distinct **polarity**, the two ends have different tubulin surfaces exposed, and polymerization and depolymerization can occur at different rates at the two ends. As a consequence, microtubules often grow at one end and disassemble at the other. Such **"treadmilling"** may be important in movement of chromosomes in neuronal migration[322] and in fast axonal transport of macromolecules (Chapter 30).[323] During mitosis the minus ends of the microtubules are believed to be tightly anchored at the centrosome while subunit exchanges can occur at the plus ends[323,324] where the β subunits are exposed. Using a phage display system (see Fig. 3-16) it could be shown that the N termini of the α subunits are exposed at the minus ends.[325] Kinesin can bind to the β subunits all along the microtubule.[326] Microtubules are formed by growth from microtubule nucleation sites in **microtubule organizing centers** found in centrosomes, spindle poles, and other locations.[327] Several proteins, including **γ-tubulin**, are required.[317,328,329] A proposed assembly pathway is illustrated in Fig. 7-34.

Isolated microtubules always contain small amounts of larger ~300-kDa **microtubule-associated proteins** (MAPS).[330] These elongated molecules may in part lie in the grooves between the tubulin subunits and in part be extended outward to form a low-density layer around the tubule.[283,309] Nerve cells that contain stable microtubules have associated stabilizing proteins.[331] A family of proteins formed by differential splicing of mRNA are known as **tau**. The tau proteins are prominent components of the cytoskeleton of neurons. They not only interact with microtubules but also undergo reversible phosphorylation. Hyperphosphorylated tau is the primary component of the paired helical filaments found in the brains of persons with Alzheimer disease.[330]

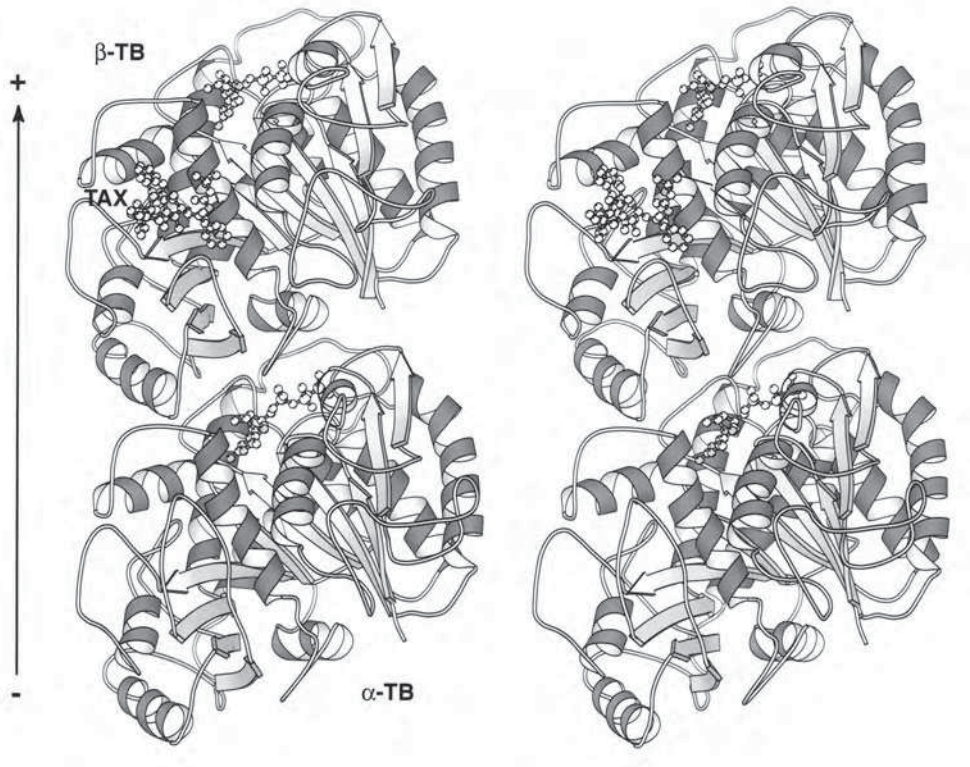



**Figure 7-33** Stereoscopic ribbon diagram of the tubulin dimer with α-tubulin with bound GTP at the top and β-tubulin with bound GDP at the bottom. The β-tubulin subunit also contains a bound molecule of taxotere (see Box 7-D) which is labeled TAX. This model is based upon electron crystallography of zinc-induced tubulin sheets at 0.37-nm resolution and is thought to approximate closely the packing of the tubulin monomers in microtubules.[315] The arrow at the left points toward the plus end of the microtubule. Courtesy of Kenneth H. Downing.

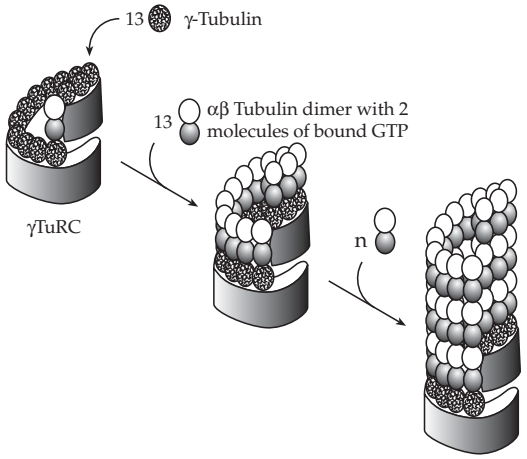



**Figure 7-34** Growth of a microtubule from a γ-tubulin ring complex (γTURC). The helical γ-tubulin rings are formed in the microtubule organizing centers which, in animal cells, are the centrosomes. Thirteen γ-tubulin subunits are shown in a hypothetical array formed together with a base of other molecules of unknown structure. The microtubule grows by addition of successive layers of α/β-tubulin dimers, each a split ring of 13 dimers with the β-tubulin subunits toward the base, the **negative end**, and the α-tubulin subunits toward the growing **positive end**. After Zheng *et al*.[317]

## References


1. Williams, D. H., Searle, M. S., Mackay, J. P., Gerhard, U., and Maplestone, R. A. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1172–1178
2. Weber, G. (1992) *Protein Interactions*, Chapman and Hall, London
3. Scatchard, G. (1949) *Ann. N.Y. Acad. Sci.* **51**, 660–672
4. Zierler, K. (1989) *Trends Biochem. Sci.* **14**, 314–317
5. Munson, P. J., Rodbard, D., and Klotz, I. M. (1983) *Science* **220**, 979–981
6. Deranleau, D. A. (1969) *J. Am. Chem. Soc.* **91**, 4044–4049
7. Dowd, J. E., and Riggs, D. S. (1965) *J. Biol. Chem.* **240**, 863–869
8. Feldman, H. A. (1983) *J. Biol. Chem.* **258**, 12865–12867

8a. Dunford, H. B. (1984) *J. Chem. Educ.* **61**, 129–132

8b. Klotz, I. M. (1997) *Ligand-Receptor Energetics*, Wiley, New York

9. Conners, K. A. (1987) *Binding Constants*, Wiley, New York
10. Motulsky, H. J., and Ransnas, L. A. (1987) *FASEB J.* **1**, 365–374
11. Schejter, A., and Margoliash, E. (1985) *Trends Biochem. Sci.* **10**, 490–492
12. Steinhardt, J., and Reynolds, J. A. (1969) *Multiple Equilibria in Proteins*, Academic Press, New York
13. Koshland, D. E., Jr. (1970) in *The Enzymes*, 3rd ed., Vol. 1 (Boyer, P. D., ed), pp. 341–396, Academic Press, New York
14. Tanford, C. (1961) *Physical Chemistry of Macromolecules*, Wiley, New York (pp. 533–534)
15. Edsall, J. T., and Wyman, J. (1958) *Biophysical Chemistry*, Academic Press, New York
16. Westheimer, F. H., and Shookhoff, M. W. (1939) *J. Am. Chem. Soc.* **61**, 555–560
17. Honig, B., and Nicholls, A. (1995) *Science* **268**, 1144–1149
18. Rajasekaran, E., Jayaram, B., and Honig, B. (1994) *J. Am. Chem. Soc.* **116**, 8238–8240
19. Tanford, C. (1961) *Physical Chemistry of Macromolecules*, Wiley, New York (Chapter 8)
20. Maier, G. D., and Metzler, D. E. (1957) *J. Am. Chem. Soc.* **79**, 4386–4391
21. Metzler, D. E. (1960) in *The Enzymes*, 2nd ed., Vol. 2 (Boyer, P. D., Lardy, H., and Myrback, K., eds), pp. 295–337, Academic Press, New York
22. Applequist, J. (1977) *J. Chem. Educ.* **54**, 417–419
23. Hill, T. (1962) *Introduction to Statistical Thermodynamics*, Addison–Wesley, Reading, Massachusetts (pp. 235–241)
24. Weiss, J. N. (1997) *FASEB J.* **11**, 835–841
25. Wang, X.-G., and Engel, P. C. (1995) *Biochemistry* **34**, 11417–11422
26. Damle, V. N. (1972) *Biopolymers* **11**, 1789–816
27. Watson, J. D. (1976) *Molecular Biology of the Gene*, 3rd ed., Benjamin, Menlo Park, California (p. 98)
28. Bernal, J. D. (1967) *J. Mol. Biol.* **24**, 379–390
29. Blundell, T. L., and Johnson, L. N. (1976) *Protein Crystallography*, Academic Press, New York
30. Monod, J., Wyman, J., and Changeux, J. D. (1965) *J. Mol. Biol.* **12**, 88–118

30a. Dill, K. A. (1997) *J. Biol. Chem.* **272**, 701–704

30b. Boresch, S., and Karplus, M. (1995) *J. Mol. Biol.* **254**, 801–807

30c. Tamura, A., and Privalov, P. L. (1997) *J. Mol. Biol.* **273**, 1048–1060

31. Zhang, R.-G., Westbrook, M. L., Westbrook, E. M., Scott, D. L., Otwinowski, Z., Maulik, P. R., Reed, R. A., and Shipley, G. G. (1995) *J. Mol. Biol.* **251**, 550–562

31a. Marvin, D. A. (1998) *Current Opinion in Structural Biology* **8**, 150–158

32. Marvin, D. A., Hale, R. D., Nave, C., and Citterich, M. H. (1994) *J. Mol. Biol.* **235**, 260–286
33. Glucksman, M. J., Bhattacharjee, S., and Makowski, L. (1992) *J. Mol. Biol.* **226**, 455–470
34. Day, L. A., Marzec, C. J., Reisberg, S. A., and Casadevall, A. (1988) *Ann. Rev. Biophys. Biophys. Chem.* **17**, 509–539
35. Williams, K. A., Glibowicka, M., Li, Z., Li, H., Khan, A. R., Chen, Y. M. Y., Wang, J., Marvin, D. A., and Deber, C. M. (1995) *J. Mol. Biol.* **252**, 6–14
36. Overman, S. A., Tsuboi, M., and Thomas, G. J., Jr. (1996) *J. Mol. Biol.* **259**, 331–336
37. Makowski, L. (1992) *J. Mol. Biol.* **228**, 885–892
38. Rakonjac, J., and Model, P. (1998) *J. Mol. Biol.* **282**, 25–41

38a. Welsh, L. C., Marvin, D. A., and Perham, R. N. (1998) *J. Mol. Biol.* **284**, 1265–1271

39. Liu, D. J., and Day, L. A. (1994) *Science* **265**, 671–674
40. Bhyravbhatla, B., Watowich, S. J., and Caspar, D. L. D. (1998) *Biophys. J.* **74**, 604–615
41. Fraenkel-Conrat, H. (1964) *Sci. Am.* **211**(Oct), 47–54
42. Butler, P. J. G., and Klug, A. (1978) *Sci. Am.* **239**(Nov), 62–69
43. Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R., and Klug, A. (1978) *Nature (London)* **276**, 362–368
44. Holmes, K. C. (1980) *Trends Biochem. Sci.* **5**, 4–7

44a. Eisenstein, M., Shariv, I., Koren, G., Friesem, A. A., and Katchalski-Katzir, E. (1997) *J. Mol. Biol.* **266**, 135–143

45. Raghavendra, K., Kelly, J. A., Khairallah, L., and Schuster, T. M. (1988) *Biochemistry* **27**, 7583–7588
46. Wang, H., and Stubbs, G. (1994) *J. Mol. Biol.* **239**, 371–384
47. Namba, K., Casper, D. L. D., and Stubbs, G. J. (1985) *Science* **227**, 773–776
48. Eisenstein, B. I. (1987) Escherichia coli *and* Salmonella typhimurium, Am. Soc. Microbiology, Washington, D.C., FC Niedhardt, ed. (pp. 84–90)
49. Bullitt, E., and Makowski, L. (1995) *Nature (London)* **373**, 164–167
50. Bullitt, E., Jones, C. H., Striker, R., Soto, G., Jacob-Dubuisson, F., Pinkner, J., Wick, M. J., Makowski, L., and Hultgren, S. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12890–12895
51. Bullitt, E., and Makowski, L. (1998) *Biophys. J.* **74**, 623–632
52. Gong, M., and Makowski, L. (1992) *J. Mol. Biol.* **228**, 735–742
53. Kuehn, M. J., Ogg, D. J., Kihlberg, J., Slonim, L. N., Flemmer, K., Bergfors, T., and Hultgren, S. J. (1993) *Science* **262**, 1234–1241

53a. Sauer, F. G., Fütterer, K., Pinkner, J. S., Dodson, K. W., Hultgren, S. J., and Waksman, G. (1999) *Science* **285**, 1058–1061

53b. Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J., and Knight, S. D. (1999) *Science* **285**, 1061–1065

53c. Skerker, J. M., and Shapiro, L. (2000) *EMBO J.* **19**, 3223–3234

53d. Hazes, B., Sastry, P. A., Hayakawa, K., Read, R. J., and Irvin, R. T. (2000) *J. Mol. Biol.* **299**, 1005–1017

53e. Collinson, S. K., Parker, J. M. R., Hodges, R. S., and Kay, W. W. (1999) *J. Mol. Biol.* **290**, 741–756

53f. Murphy, F. V., IV, Sweet, R. M., and Churchill, M. E. A. (1999) *EMBO J.* **18**, 6610–6618

54. Hanson, M. S., and Brinton, C. C., Jr. (1988) *Nature (London)* **332**, 265–268
55. St. Geme, J. W., III, Pinkner, J. S., Krasan, G. P., Heuser, J., and Bullitt, E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 11913–11918
56. Parge, H. E., Bernstein, S. L., Deal, C. D., McRee, D. E., Christensen, D., Capozza, M. A., Kays, B. W., Fieser, T. M., Draper, D., So, M., Getzoff, E. D., and Tainer, J. A. (1990) *J. Biol. Chem.* **265**, 2278–2285

## References


57. Parge, H. E., Forest, K. T., Hickey, M. J., Christensen, D. A., Getzoff, E. D., and Tainer, J. A. (1995) *Nature (London)* **378**, 32–38
58. Folkhard, W., Leonard, K. R., Malsey, S., Marvin, D. A., Dubochet, J., Engel, A., Achtman, M., and Helmuth, R. (1979) *J. Mol. Biol.* **130**, 145–160
59. Paiva, W. D., Grossman, T., and Silverman, P. M. (1992) *J. Biol. Chem.* **267**, 26191–26197
60. Rubenstein, P. A. (1990) *BioEssays* **12**, 309–315
61. Venkatesh, B., Tay, B. H., Elgar, G., and Brenner, S. (1996) *J. Mol. Biol.* **259**, 655–665
62. Holmes, K. C., Popp, D., Gebhard, W., and Kabsch, W. (1990) *Nature (London)* **347**, 44–49
63. Lorenz, M., Poole, K. J. V., Popp, D., Rosenbaum, G., and Holmes, K. C. (1995) *J. Mol. Biol.* **246**, 108–119
64. Valentine, R. C. (1969) in *Symmetry and Function of the Biological Systems at the Macromolecular Level* (Engstrom, A., and Strandberg, B., eds), p. 165, Wiley, New York (11th Nobel Symp.)
65. Klotz, I. M., Darnall, D. W., and Langerman, N. R. (1975) in *The Proteins*, 3rd ed., Vol. 1 (Neurath, H., and Hill, R. L., eds), p. 293, Academic Press, New York
66. Momany, C., Ernst, S., Ghosh, R., Chang, N.-L., and Hackert, M. L. (1995) *J. Mol. Biol.* **252**, 643–655
67. Almassy, R. J., Janson, C. A., Hamlin, R., Xuong, N.-H., and Eisenberg, D. (1986) *Nature (London)* **323**, 304–309
68. Harrison, S. C. (1984) *Trends Biochem. Sci.* **9**, 345–351
69. Rossmann, M. G., and Johnson, J. E. (1989) *Ann. Rev. Biochem.* **58**, 533–573
70. Johnson, J. E., and Speir, J. A. (1997) *J. Mol. Biol.* **269**, 665–675
71. Watson, J. D. (1976) *Molecular Biology of the Gene*, 3rd ed., Benjamin, Menlo Park, California (p. 107)
72. Unge, T., Liljas, L., Strandberg, B., Vaara, I., Kannan, K. K., Fridborg, K., Nordman, C. E., and Lentz, P. J., Jr. (1980) *Nature (London)* **285**, 373–377
73. Küpper, H., Keller, W., Kurz, C., Forss, S., Schaller, H., Franze, R., Stohmaier, K., Marquardt, O., Zaslavsky, V. G., and Hofschneider, P. H. (1981) *Nature (London)* **289**, 555–559
74. Trikha, J., Theil, E. C., and Allewell, N. M. (1995) *J. Mol. Biol.* **248**, 949–967
75. Larson, S. B., Day, J., Greenwood, A., and McPherson, A. (1998) *J. Mol. Biol.* **277**, 37–59
76. Hogle, J. M., Chow, M., and Filman, D. J. (1985) *Science* **229**, 1358–1365
77. Hogle, J. M., Chow, M., and Filman, D. J. (1987) *Sci. Am.* **256** (Mar), 42–49
78. Olson, N. H., Kolatkar, P. R., Oliveira, M. A., Cheng, R. H., Greve, J. M., McClelland, A., Baker, T. S., and Rossmann, M. G. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 507–511
79. Hadfield, A. T., Oliveira, M. A., Kim, K. H., Minor, I., Kremer, M. J., Heinz, B. A., Shepard, D., Pevear, D. C., Rueckert, R. R., and Rossmann, M. G. (1995) *J. Mol. Biol.* **253**, 61–73
80. Wu, H., and Rossmann, M. G. (1993) *J. Mol. Biol.* **233**, 231–244
81. Luo, M., Vriend, G., Kamer, G., Minor, I., Arnold, E., Rossmann, M. G., Boege, U., Scraba, D. G., Duke, G. M., and Palmenberg, A. C. (1987) *Science* **235**, 182–191
82. Tikhonenko, A. S. (1970) *Ultrastructure of Bacterial Viruses*, Plenum, New York
83. Birktoft, J. J., Rhodes, G., and Banaszak, L. J. (1989) *Biochemistry* **28**, 6065–6081
84. Skarzynski, T., Moody, P. C. E., and Wonacott, A. J. (1987) *J. Mol. Biol.* **193**, 171–187
85. Zimmerle, C. T., and Alter, G. M. (1993) *Biochemistry* **32**, 12743–12748
86. Hodgkin, D. C. (1975) *Nature (London)* **255**, 103
87. Whittingham, J. L., Chaudhuri, S., Dodson, E. J., Moody, P. C. E., and Dodson, G. G. (1995) *Biochemistry* **34**, 15553–15563
88. Chang, X., Jorgensen, A. M. M., Bardrum, P., and Led, J. J. (1997) *Biochemistry* **36**, 9409–9422
89. Golmohammadi, R., Valegård, K., Fridborg, K., and Liljas, L. (1993) *J. Mol. Biol.* **234**, 620–639
90. Liljas, L., Fridborg, K., Valegård, K., Bundule, M., and Pumpens, P. (1994) *J. Mol. Biol.* **244**, 279–290
91. Tars, K., Bundule, M., Fridborg, K., and Lijas, L. (1997) *J. Mol. Biol.* **271**, 759–773
92. Abad-Zapatero, C., Abdel-Meguid, S. S., Johnson, J. E., Leslie, A. G. W., Rayment, I., Rossmann, M. G., Suck, D., and Tsukihara, T. (1980) *Nature (London)* **286**, 33–39
93. Athappilly, F. K., Murali, R., Rux, J. J., Cai, Z., and Burnett, R. M. (1994) *J. Mol. Biol.* **242**, 430–455
94. Stewart, P. L., Chiu, C. Y., Huang, S., Muir, T., Zhao, Y., Chait, B., Mathias, P., and Nemerow, G. R. (1997) *EMBO J.* **16**, 1189–1198
95. Caspar, D. L. D., and Klug, A. (1962) *Cold Spring Harbor Symposia on Quantitative Biology* **27**, 1–24
96. Johnson, J. E. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 27–33
97. Steven, A. C., Trus, B. L., Booy, F. P., Cheng, N., Zlotnick, A., Caston, J. R., and Conway, J. F. (1997) *FASEB J.* **11**, 733–742
98. Rayment, I., Baker, T. S., Caspar, D. L. D., and Murakami, W. T. (1982) *Nature (London)* **295**, 110–115
99. Liddington, R. C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T. L., and Harrison, S. C. (1991) *Nature (London)* **354**, 278–284
100. Champness, J. N., Bloomer, A. C., Bricogne, G., Butler, P. J. G., and Klug, A. (1976) *Nature (London)* **259**, 20–24
101. Fukuda, M., and Okada, Y. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4035–4038
102. McDonald, R. C., Engelman, D. M., and Steita, T. A. (1979) *J. Biol. Chem.* **254**, 2942–2943
103. Galloway, J. (1984) *Trends Biochem. Sci.* **9**, 233–238
104. Kantrowitz, E. R., and Lipscomb, W. N. (1990) *Trends Biochem. Sci.* **15**, 53–59
105. Zhou, B.-B., and Schachman, H. K. (1993) *Protein Sci.* **2**, 103–112
106. Xi, X. G., De Staercke, C., Van Vliet, F., Triniolles, F., Jacobs, A., Stas, P. P., Ladjimi, M. M., Simon, V., Cunin, R., and Hervé, G. (1994) *J. Mol. Biol.* **242**, 139–149
107. Fetler, L., Vachette, P., Hervé, G., and Ladjimi, M. M. (1995) *Biochemistry* **34**, 15654–15660
108. Honzatko, R. B., Crawford, J. L., Monaco, H. L., Ladner, J. E., Edwards, B. F. D., Evans, D. R., Warren, S. G., Wiley, D. C., Ladner, R. C., and Lipscomb, W. N. (1982) *J. Mol. Biol.* **160**, 219–263
109. Krause, K. L., Volz, K. W., and Lipscomb, W. N. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1643–1647
110. Thomas, A., Field, M. J., and Perahia, D. (1996) *J. Mol. Biol.* **261**, 490–506

110a. Brunori, M. (1999) *Trends Biochem. Sci.* **24**, 158–161

110b. Di Cera, E., ed. (1998) *Advances in Protein Chemistry, Linkage Thermodynamics of Macromolecular Interactions*, Vol. 51, Academic Press, San Diego, California

111. Bloom, C. R., Choi, W. E., Brzovic, P. S., Ha, J. J., Huang, S.-T., Kaarsholm, N. C., and Dunn, M. F. (1995) *J. Mol. Biol.* **245**, 324–330
112. Hammes, G. G., and Wu, C.-W. (1971) *Science* **172**, 1205–1211
113. Koshland, D. E., Jr., Némethy, G., and Filmer, D. (1966) *Biochemistry* **5**, 365–385
114. Cornish-Bowden, A. J., and Koshland, D. E., Jr. (1970) *J. Biol. Chem.* **245**, 6241–6250
115. Cornish-Bowden, A. J., and Koshland, D. E., Jr. (1971) *J. Biol. Chem.* **246**, 3092–3102
116. Darnall, D. W., and Klotz, I. M. (1972) *Arch. Biochem. Biophys.* **149**, 1–14
117. Doci, Y., Sugita, Y., and Yoneyama, Y. (1973) *J. Biol. Chem.* **248**, 2354–2363
118. Perutz, M. F., Muirhead, H., Cox, J. M., and Goaman, L. C. G. (1968) *Nature (London)* **219**, 131–139
119. Dickerson, R. E., and Geis, I. (1983) *Hemoglobin: Structure, Function, Evolution and Pathology*, Benjamin / Cummings Publ., Redwood City, California
120. Antonini, E., and Brunori, M. (1971) *Hemoglobin and Myoglobin in Their Reactions with Ligands*, North-Hollard, Publ., Amsterdam
121. Kendrew, J. C. (1961) *Sci. Am.* **205**(Dec), 96–110
122. Fermi, G., Perutz, M. F., Shaanan, B., and Fourme, R. (1984) *J. Mol. Biol.* **175**, 159–174
123. Benesch, R., and Benesch, R. E. (1974) *Science* **185**, 905–908
124. Borgstahl, G. E. O., Rogers, P. H., and Arnone, A. (1994) *J. Mol. Biol.* **236**, 831–843
125. Imai, K. (1979) *J. Mol. Biol.* **133**, 233–247
126. Amiconi, G., Antonini, E., Brunori, M., Wyman, J., and Zolla, L. (1981) *J. Mol. Biol.* **152**, 111–129
127. Flanagan, M. A., Ackers, G. K., Matthew, J. B., Hanania, G. I. H., and Gurd, F. R. N. (1981) *Biochemistry* **20**, 7439–7449

127a. Shibayama, N. (1999) *J. Mol. Biol.* **285**, 1383–1388

128. Waxman, L. (1971) *J. Biol. Chem.* **246**, 7318–7327
129. Doyle, M. L., Di Cera, E., and Gill, S. J. (1988) *Biochemistry* **27**, 820–824
130. Edelstein, S. J. (1996) *J. Mol. Biol.* **257**, 737–744
131. Imai, K. (1983) *J. Mol. Biol.* **167**, 741–749

131a. Gibson, Q. H. (1999) *Biochemistry* **38**, 5191–5199

132. Paoli, M., Liddington, R., Tame, J., Wilkinson, A., and Dodson, G. (1996) *J. Mol. Biol.* **256**, 775–792
133. Henry, E. R., Jones, C. M., Hofrichter, J., and Eaton, W. A. (1997) *Biochemistry* **36**, 6511–6528
134. Silva, M. M., Rogers, P. H., and Arnone, A. (1992) *J. Biol. Chem.* **267**, 17248–17256
135. Jayaraman, V., and Spiro, T. G. (1995) *Biochemistry* **34**, 4511–4515
136. Schumacher, M. A., Zheleznova, E. E., Poundstone, K. S., Kluger, R., Jones, R. T., and Brennan, R. G. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7841–7844

136a. Tame, J. R. H. (1999) *Trends Biochem. Sci.* **24**, 372–377

137. Manning, L. R., Jenkins, W. T., Hess, J. R., Vandegriff, K., Winslow, R. M., and Manning, J. M. (1996) *Protein Sci.* **5**, 775–781
138. Holt, J. M., and Ackers, G. K. (1995) *FASEB J.* **9**, 210–218
139. Perutz, M. F., Shih, D. T.-b, and Williamson, D. (1994) *J. Mol. Biol.* **239**, 555–560
140. Perutz, M. F., Muirhead, H., Cox, J. M., Goaman, L. C. G., Mathews, F. S., McGandy, L. E., and Webb, L. E. (1968) *Nature (London)* **219**, 29–32
141. Perutz, M. (1990) *Mechanisms of Cooperativity and Allosteric Regulation in Proteins*, Cambridge Univ. Press, London
142. O'Donnell, S., Mandaro, R., Schuster, T. M., and Arnone, A. (1979) *J. Biol. Chem.* **254**, 12204–12208
143. Perutz, M. F. (1970) *Nature (London)* **228**, 726–739
144. Dickerson, R. E. (1972) *Ann. Rev. Biochem.* **41**, 815–842

## References


145. Ishimori, K., Hashimoto, M., Imai, K., Fushitani, K., Miyazaki, G., Morimoto, H., Wada, Y., and Morishima, I. (1994) *Biochemistry* **33**, 2546–2553
146. Chao, H., Sönnichsen, F. D., DeLuca, C. I., Sykes, B. D., and Davies, P. L. (1994) *Protein Sci.* **3**, 1760–1769
147. Kim, H.-W., Shen, T.-J., Ho, N. T., Zou, M., Tam, M. F., and Ho, C. (1996) *Biochemistry* **35**, 6620–6627
148. Kavanaugh, J. S., Weydert, J. A., Rogers, P. H., and Arnone, A. (1998) *Biochemistry* **37**, 4358–4373
149. Kiger, L., Klinger, A. L., Kwiatkowski, L. D., De Young, A., Doyle, M. L., Holt, J. M., Noble, R. W., and Ackers, G. K. (1998) *Biochemistry* **37**, 4336–4345
150. Hoard, J. L., and Scheidt, W. R. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3919–3922
151. Fermi, G., Perutz, M. F., and Shulman, R. G. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6167–6168
152. De Baere, I., Perutz, M. F., Kiger, L., Marden, M. C., and Poyart, C. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1594–1597
153. Bohr, C., Hasselbalch, K., and Krogh, A. (1904) *Skand. Arch. Physiol.* **16**, 402–
154. Wyman, J., Jr. (1948) *Adv. Prot. Chem.* **4**, 407–531
155. Benesch, R. E., and Benesch, R. (1974) *Adv. Prot. Chem.* **28**, 211–237
156. Fang, T.-Y., Zou, M., Simplaceanu, V., Ho, N. T., and Ho, C. (1999) *Biochemistry* **38**, 13423–13432
157. Fronticelli, C., Pechik, I., Brinigar, W. S., Kowalczyk, J., and Gilliland, G. L. (1994) *J. Biol. Chem.* **269**, 23965–23969
158. Perrella, M., Ripamonti, M., and Caccia, S. (1998) *Biochemistry* **37**, 2017–2028
159. Arnone, A. (1972) *Nature (London)* **237**, 146–149
160. Benesch, R., and Benesch, R. E. (1969) *Nature (London)* **221**, 618–622
161. Brewer, G. J., and Eaton, J. W. (1971) *Science* **171**, 1205–1211
162. Perutz, M. F., Bauer, C., Gros, G., Leclercq, F., Vandecasserie, C., Schnek, A. G., Braunitzer, G., Friday, A. E., and Joysey, K. A. (1981) *Nature (London)* **291**, 682–684
163. Komiyama, N. H., Miyazaki, G., Tame, J., and Nagai, K. (1995) *Nature (London)* **373**, 244–246
164. Fantl, W. J., Di Donato, A., Manning, J. M., Rogers, P. H., and Arnone, A. (1987) *J. Biol. Chem.* **87**, 12700–12713
165. Sage, J. T., and Jee, W. (1997) *J. Mol. Biol.* **274**, 21–26
166. Teng, T.-Y., Srajer, V., and Moffat, K. (1997) *Biochemistry* **36**, 12087–12100
166a. Ostermann, A., Waschipky, R., Parak, F. G., and Nienhaus, G. U. (2000) *Nature (London)* **404**, 205–208
167. Chance, M. R., Miller, L. M., Fischetti, R. F., Scheuring, E., Huang, W.-X., Sclavi, B., Hai, Y., and Sullivan, M. (1996) *Biochemistry* **35**, 9014–9023
168. Ösapay, K., Theriault, Y., Wright, P. E., and Case, D. A. (1994) *J. Mol. Biol.* **244**, 183–197
169. Martin, K. D., and Parkhurst, L. J. (1990) *Biochemistry* **29**, 5718–5726
170. Tetreau, C., Lavalette, D., Momenteau, M., Fischer, J., and Weiss, R. (1994) *J. Am. Chem. Soc.* **116**, 11840–11848
171. Huang, Y., and Ackers, G. K. (1995) *Biochemistry* **34**, 6316–6327
172. Paoli, M., Dodson, G., Liddington, R. C., and Wilkinson, A. J. (1997) *J. Mol. Biol.* **271**, 161–167
173. Daugherty, M. A., Shea, M. A., and Ackers, G. K. (1994) *Biochemistry* **33**, 10345–10357
174. Stamler, J. S., Jia, L., Eu, J. P., McMahon, T. J., Demchenko, I. T., Bonaventura, J., Gernert, K., and Piantadosi, C. A. (1997) *Science* **276**, 2034–2037
175. Upmacis, R. K., Hajjar, D. P., Chait, B. T., and Mirza, U. A. (1997) *J. Am. Chem. Soc.* **119**, 10424–10429
176. Miller, L. M., Pedraza, A. J., and Chance, M. R. (1997) *Biochemistry* **36**, 12199–12207
177. Gow, A. J., and Stamler, J. S. (1998) *Nature (London)* **391**, 169–173
178. Lutter, L. C., Halvorson, H. R., and Calladine, C. R. (1996) *J. Mol. Biol.* **261**, 620–633
179. Jia, L., Bonaventura, C., Bonaventura, J., and Stamler, J. S. (1996) *Nature (London)* **380**, 221–226
180. Rizzi, M., Wittenberg, J. B., Coda, A., Fasano, M., Ascenzi, P., and Bolognesi, M. (1994) *J. Mol. Biol.* **244**, 86–99
181. Rizzi, M., Wittenberg, J. B., Coda, A., Ascenzi, P., and Bolognesi, M. (1996) *J. Mol. Biol.* **258**, 1–5
182. Nguyen, B. D., Zhao, X., Vyas, K., La Mar, G. N., Lile, R. A., Brucker, E. A., Phillips, G. N., Jr., Olson, J. S., and Wittenberg, J. B. (1998) *J. Biol. Chem.* **273**, 9517–9526
183. Weatherall, D. J., Clegg, J. B., Higgs, D. R., and Wood, W. G. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 1 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 3417–3484, McGraw-Hill, New York
184. Bank, A., Mears, J. G., and Ramirez, F. (1980) *Science* **207**, 486–493
185. Hirsch, R. E., Juszczak, L. J., Fataliev, N. A., Friedman, J. M., and Nagel, R. L. (1999) *J. Biol. Chem.* **274**, 13777–13782
186. Arnone, A., Thillet, J., and Rosa, J. (1981) *J. Biol. Chem.* **256**, 8545–8552
187. Lehmann, H., and Huntsman, R. G. (1966) *Man's Haemoglobins*, North-Holland Publ. Co., Amsterdam (p. 246)
188. Olson, J. S., Gibson, Q. H., Nagel, R. L., and Hamilton, H. B. (1972) *J. Biol. Chem.* **247**, 7485–7493
189. Poyart, C., Bursaux, E., Arnone, A., Bonaventura, J., and Bonaventura, C. (1980) *J. Biol. Chem.* **255**, 9465–9473
189a. Burmester, T., Weich, B., Reinhardt, S., and Hankeln, T. (2000) *Nature* **407**, 520–523
190. Tomita, S. (1981) *J. Biol. Chem.* **256**, 9495–9500
191. Giardina, B., Scatena, R., Clementi, M. E., Cerroni, L., Nuutinen, M., Brix, O., Sletten, S. N., Castagnola, M., and Condò, S. G. (1993) *J. Mol. Biol.* **229**, 512–516
192. Clementi, M. E., Scatena, R., Mordente, A., Condò, S. G., Castagnola, M., and Giardina, B. (1996) *J. Mol. Biol.* **255**, 229–234
192a. Burr, A. H. J., Hunt, P., Wagar, D. R., Dewilde, S., Blaxter, M. L., Vanfleteren, J. R., and Moens, L. (2000) *J. Biol. Chem.* **275**, 4810–4815
193. Lee, H. C., Wittenberg, J. B., and Peisach, J. (1993) *Biochemistry* **32**, 11500–11506
194. Harutyunyan, E. H., Safonova, T. N., Kuranova, I. P., Popov, A. N., Teplyakov, A. V., Obmolova, G. V., Rusakov, A. A., Vainshtein, B. K., Dodson, G. G., Wilson, J. C., and Perutz, M. F. (1995) *J. Mol. Biol.* **251**, 104–115
195. Hargrove, M. S., Barry, J. K., Brucker, E. A., Berry, M. B., Phillips, G. N., Jr., Olson, J. S., Arredondo-Peter, R., Dean, J. M., Klucas, R. V., and Sarath, G. (1997) *J. Mol. Biol.* **266**, 1032–1042
196. Trevaskis, B., Watts, R. A., Andersson, C. R., Llewellyn, D. J., Hargrove, M. S., Olson, J. S., Dennis, E. S., and Peacock, W. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12230–12234
196a. Hargrove, M. S., Brucker, E. A., Stec, B., Sarath, G., Arredondo-Peter, R., Klucas, R. V., Olson, J. S., and Phillips, G. N., Jr. (2000) *Structure* **8**, 1005–1014
197. Hardison, R. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5675–5679
198. Potts, M., Angeloni, S. V., Ebel, R. E., and Bassam, D. (1992) *Science* **256**, 1690–1692
199. Yeh, S.-R., Couture, M., Ouellet, Y., Guertin, M., and Rousseau, D. L. (2000) *J. Biol. Chem.* **275**, 1679–1684
200. Aronson, H.-E. G., Royer, W. E., Jr., and Hendrickson, W. A. (1994) *Protein Sci.* **3**, 1706–1711
201. Honzatko, R. B., Hendrickson, W. A., and Love, W. E. (1985) *J. Mol. Biol.* **184**, 147–164
201a. Qiu, Y., Maillett, D. H., Knapp, J., Olson, J. S., and Riggs, A. F. (2000) *J. Biol. Chem.* **275**, 13517–13528
202. Condon, P. J., and Royer, W. E. J. (1994) *J. Biol. Chem.* **269**, 25259–25267
203. Royer, W. E., Jr., Heard, K. S., Harrington, D. J., and Chiancone, E. (1995) *J. Mol. Biol.* **253**, 168–186
204. Mozzarelli, A., Bettati, S., Rivetti, C., Rossi, G. L., Colotti, G., and Chiancone, E. (1996) *J. Biol. Chem.* **271**, 3627–3632
205. Yang, J., Kloek, A. P., Goldberg, D. E., and Mathews, F. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4224–4228
206. Huang, S., Huang, J., Kloek, A. P., Goldberg, D. E., and Friedman, J. M. (1996) *J. Biol. Chem.* **271**, 958–962
207. Lamy, J., Kuchumov, A., Taveau, J.-C., Vinogradov, S. N., and Lamy, J. N. (2000) *J. Mol. Biol.* **298**, 633–647
208. de Haas, F., Taveau, J.-C., Boisset, N., Lambert, O., Vinogradov, S. N., and Lamy, J. N. (1996) *J. Mol. Biol.* **255**, 140–153
209. de Haas, F., Zal, F., You, V., Lallier, F., Toulmond, A., and Lamy, J. N. (1996) *J. Mol. Biol.* **264**, 111–120
210. de Haas, F., Boisset, N., Taveau, J.-C., Lambert, O., Vinogradov, S. N., and Lamy, J. N. (1996) *Biophys. J.* **70**, 1973–1984
211. Zuckerkandl, E. (1965) *Sci. Am.* **212**(May), 110–118
212. Perutz, M. F. (1986) *Nature (London)* **322**, 405
213. Wood, W. B., and Edgar, R. S. (1967) *Sci. Am.* **217**(Jul), 60–74
214. Kikuchi, Y., and King, J. (1975) *J. Mol. Biol.* **99**, 645–672
215. Casjens, S., and King, J. (1975) *Ann. Rev. Biochem.* **44**, 555–611
216. Pollard, T. D., and Craig, S. W. (1982) *Trends Biochem. Sci.* **7**, 55–58
217. Ito, K., Date, T., and Wickner, W. (1980) *J. Biol. Chem.* **255**, 2123–2130
218. McDonnell, P. A., Shon, K., Kim, Y., and Opella, S. J. (1993) *J. Mol. Biol.* **233**, 447–463
219. Sanders, J. C., Haris, P. I., Chapman, D., Otto, C., and Hemminga, M. A. (1993) *Biochemistry* **32**, 12446–12454
220. Guy-Caffey, J. K., and Webster, R. E. (1993) *J. Biol. Chem.* **268**, 5496–5503
221. Duda, R. L., Martincic, K., and Hendrix, R. W. (1995) *J. Mol. Biol.* **247**, 636–647
222. Xie, Z., and Hendrix, R. W. (1995) *J. Mol. Biol.* **253**, 74–85
223. Marvik, C. J., Dokland, T., Nokling, R. H., Jacobsen, E., Larsen, T., and Lindqvist, B. H. (1995) *J. Mol. Biol.* **251**, 59–75
224. LeCuyer, K. A., Behlen, L. S., and Uhlenbeck, O. C. (1995) *Biochemistry* **34**, 10600–10606
225. McKenna, R., Ilag, L. L., and Rossmann, M. G. (1994) *J. Mol. Biol.* **237**, 517–543
226. Dokland, T., McKenna, R., Ilag, L. L., Bowman, B. R., Incardona, N. L., Fane, B. A., and Rossmann, M. G. (1997) *Nature (London)* **389**, 308–313
227. Thuman-Commike, P. A., Greene, B., Jakana, J., Prasad, B. V. V., King, J., Prevelige, P. E., Jr., and Chiu, W. (1996) *J. Mol. Biol.* **260**, 85–98
228. Parker, M. H., Stafford, W. F., III, and Prevelige, P. E., Jr. (1997) *J. Mol. Biol.* **268**, 655–665

## References


229. Butcher, S. J., Dokland, T., Ojala, P. M., Bamford, D. H., and Fuller, S. D. (1997) *EMBO J.* **16**, 4477–4487
230. Olson, N. H., Baker, T. S., Willingmann, P., and Incardona, N. L. (1992) *J. Structural Biol.* **108**, 168–175
231. Hayashi, M., Aoyama, A., Richardson, D. L., Jr., and Hayashi, M. N. (1988) in *The Bacteriophages* (Calendar, R., ed), pp. 1–71, Plenum, New York
232. Butcher, S. J., Bamford, D. H., and Fuller, S. D. (1995) *EMBO J.* **14**, 6078–6086
233. Berger, B., Shor, P. W., Tucker-Kellogg, L., and King, J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7732–7736
234. Steven, A. C., Greenstone, H. L., Booy, F. P., Black, L. W., and Ross, P. D. (1992) *J. Biol. Chem.* **228**, 870–884
235. van Driel, R. (1980) *J. Mol. Biol.* **138**, 27–42
236. King, J., and Mykolajewycz, N. (1973) *J. Mol. Biol.* **75**, 339–358
237. Feller, S. M., Ren, R., Hanafusa, H., and Baltimore, D. (1994) *Trends Biochem. Sci.* **19**, 453–458
238. Pawson, T. (1995) *Nature (London)* **373**, 573–580
239. Koch, C. A., Anderson, D., Moran, M. F., Ellis, C., and Pawson, T. (1991) *Science* **252**, 668–674
240. Maignan, S., Guilloteau, J.-P., Fromage, N., Arnoux, B., Becquart, J., and Ducruix, A. (1995) *Science* **268**, 291–293
241. Ullner, M., Selander, M., Persson, E., Stenflo, J., Drakenberg, T., and Teleman, O. (1992) *Biochemistry* **31**, 5974–5983
242. Meininger, D. P., Hunter, M. J., and Komives, E. A. (1995) *Protein Sci.* **4**, 1683–1695
243. Bersch, B., Hernandez, J.-F., Marion, D., and Arlaud, G. J. (1998) *Biochemistry* **37**, 1204–1214
244. Rand, M. D., Lindblom, A., Carlson, J., Villoutreix, B. O., and Stenflo, J. (1997) *Protein Sci.* **6**, 2059–2071
245. Müller, K., Gombert, F. O., Manning, U., Grossmüller, F., Graff, P., Zaegel, H., Zuber, J. F., Freuler, F., Tschopp, C., and Baumann, G. (1996) *J. Biol. Chem.* **271**, 16500–16505
246. Fry, M. J., Panayotou, G., Booker, G. W., and Waterfield, M. D. (1993) *Protein Sci.* **2**, 1785–1797
247. Malek, S. N., Yang, C. H., Earnshaw, W. C., Kozak, C. A., and Desiderio, S. (1996) *J. Biol. Chem.* **271**, 6952–6962
248. Tong, L., Warren, T. C., King, J., Betageri, R., Rose, J., and Jakes, S. (1996) *J. Mol. Biol.* **256**, 601–610
249. Xu, R. X., Word, J. M., Davis, D. G., Rink, M. J., Willard, D. H., Jr., and Gampe, R. T., Jr. (1995) *Biochemistry* **34**, 2107–2121
250. Mikol, V., Baumann, G., Zurini, M. G. M., and Hommel, U. (1995) *J. Mol. Biol.* **254**, 86–95
251. Pascal, S. M., Yamazaki, T., Singer, A. U., Kay, L. E., and Forman-Kay, J. D. (1995) *Biochemistry* **34**, 11353–11362
252. Zhou, M.-M., Meadows, R. P., Logan, T. M., Yoon, H. S., Wade, W. S., Ravichandran, K. S., Burakoff, S. J., and Fesik, S. W. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7784–7788
253. Metzler, W. J., Leiting, B., Pryor, K., Mueller, L., and Farmer, B. T., II. (1996) *Biochemistry* **35**, 6201–6211
254. Viguera, A. R., Arrondo, J. L. R., Musacchio, A., Saraste, M., and Serrano, L. (1994) *Biochemistry* **33**, 10925–10933
255. Liang, J., Chen, J. K., Schreiber, S. L., and Clardy, J. (1996) *J. Mol. Biol.* **257**, 632–643
256. Lim, W. A., Richards, F. M., and Fox, R. O. (1994) *Nature (London)* **372**, 375–379
257. Guruprasad, L., Dhanaraj, V., Timm, D., Blundell, T. L., Gout, I., and Waterfield, M. D. (1995) *J. Mol. Biol.* **248**, 856–866
258. van der Geer, P., and Pawson, T. (1995) *Trends Biochem. Sci.* **20**, 277–280
259. Zhou, M.-M., Ravichandran, K. S., Olejniczak, E. T., Petros, A. M., Meadows, R. P., Sattler, M., Harlan, J. E., Wade, W. S., Burakoff, S. J., and Fesik, S. W. (1995) *Nature (London)* **378**, 584–592
260. Rameh, L. E., Arvidsson, A.-k, Carraway, K. L., III, Couvillon, A. D., Rathbun, G., Crompton, A., VanRenterghem, B., Czech, M. P., Ravichandran, K. S., Burakoff, S. J., Wang, D.-S., Chen, C.-S., and Cantley, L. C. (1997) *J. Biol. Chem.* **272**, 22059–22066
261. Hemmings, B. A. (1997) *Science* **275**, 1899
262. Blomberg, N., Baraldi, E., Nilges, M., and Saraste, M. (1999) *Trends Biochem. Sci.* **24**, 441–445
263. Koshiba, S., Kigawa, T., Kim, J.-H., Shirouzu, M., Bowtell, D., and Yokoyama, S. (1997) *J. Mol. Biol.* **269**, 579–591
264. Saras, J., and Heldin, C.-H. (1996) *Trends Biochem. Sci.* **21**, 455–458
265. Cabral, J. H. M., Petosa, C., Sutcliffe, M. J., Raza, S., Byron, O., Poy, F., Marfatia, S. M., Chishti, A. H., and Liddington, R. C. (1996) *Nature (London)* **382**, 649–652
266. Ponting, C. P. (1997) *Protein Sci.* **6**, 464–468
267. Resnick, D., Pearson, A., and Krieger, M. (1994) *Trends Biochem. Sci.* **19**, 5–8
268. Chang, Y., Mochalkin, I., McCance, S. G., Cheng, B., Tulinsky, A., and Castellino, F. J. (1998) *Biochemistry* **37**, 3258–3271
269. de Vos, A. M., Ultsch, M. H., Kelley, R. F., Padmanabhan, K., Tulinsky, A., Westbrook, M. L., and Kossiakoff, A. A. (1992) *Biochemistry* **31**, 270–279
270. De Serrano, V. S., and Castellino, F. J. (1992) *Biochemistry* **31**, 11698–11706
271. Sun, M.-F., Zhao, M., and Gailani, D. (1999) *J. Biol. Chem.* **274**, 36373–36377
272. Macias, M. J., Hyvönen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996) *Nature (London)* **382**, 646–649
273. Bogusky, M. J., Dobson, C. M., and Smith, R. A. G. (1989) *Biochemistry* **28**, 6728–6735
274. Hoffmann, W., and Hauser, F. (1993) *Trends Biochem. Sci.* **18**, 239–243
275. Lamb, J. R., Tugendreich, S., and Hieter, P. (1995) *Trends Biochem. Sci.* **20**, 257–259
276. Das, A. K., Cohen, P. T. W., and Barford, D. (1998) *EMBO J.* **17**, 1192–1199
277. Barlow, P. N., Luisi, B., Milner, A., Elliott, M., and Everett, R. (1994) *J. Mol. Biol.* **237**, 201–211
278. Lichtarge, O., Yamamoto, K. R., and Cohen, F. E. (1997) *J. Mol. Biol.* **274**, 325–337
279. Saurin, A. J., Borden, K. L. B., Boddy, M. N., and Freemont, P. S. (1996) *Trends Biochem. Sci.* **21**, 208–214
280. Pérez-Alvarado, G. C., Kosa, J. L., Louis, H. A., Beckerle, M. C., Winge, D. R., and Summers, M. F. (1996) *J. Mol. Biol.* **257**, 153–174
281. Porter, K. R., and Tucker, J. B. (1981) *Sci. Am.* **244**(Mar), 57–67
282. Weber, K., and Osborn, M. (1985) *Sci. Am.* **253**(Oct), 110–120
283. Amos, L. A., and Amos, B. W. (1991) *Molecules of the Cytoskeleton*, The Guilford Press, New York
284. Lazarides, E. (1980) *Nature (London)* **283**, 249–256
285. Fuchs, E., and Weber, K. (1994) *Ann. Rev. Biochem.* **63**, 345–382
286. Fuchs, E., and Cleveland, D. W. (1998) *Science* **279**, 514–519
287. Kaufmann, S. H. (1989) *J. Biol. Chem.* **264**, 13946–13955
288. Ip, W., Hartzer, M. K., Pang, Y.-Y. S., and Robson, R. M. (1985) *J. Mol. Biol.* **183**, 365–375
289. Herrmann, H., Häner, M., Brettel, M., Müller, S. A., Goldie, K. N., Fedtke, B., Lustig, A., Franke, W. W., and Aebi, U. (1996) *J. Mol. Biol.* **264**, 933–953

289a. Herrmann, H., Strelkov, S. V., Feja, B., Rogers, K. R., Brettel, M., Lustig, A., Häner, M., Parry, D. A. D., Steinert, P. M., Burkhard, P., and Aebi, U. (2000) *J. Mol. Biol.* **298**, 817–832

289b. Heimburg, T., Schuenemann, J., Weber, K., and Geisler, N. (1996) *Biochemistry* **35**, 1375–1382

290. Geisler, N., Schünemann, J., and Weber, K. (1992) *Eur. J. Biochem.* **206**, 841–852

290a. Bellin, R. M., Sernett, S. W., Becker, B., Ip, W., Huiatt, T. W., and Robson, R. M. (1999) *J. Biol. Chem.* **274**, 29493–29499

291. Leterrier, J. F., Käs, J., Hartwig, J., Vegners, R., and Janmey, P. A. (1996) *J. Biol. Chem.* **271**, 15687–15694
292. Betts, J. C., Blackstock, W. P., Ward, M. A., and Anderton, B. H. (1997) *J. Biol. Chem.* **272**, 12922–12927
293. Carter, J., Gragerov, A., Konvicka, K., Elder, G., Weinstein, H., and Lazzarini, R. A. (1998) *J. Biol. Chem.* **273**, 5101–5108
294. Osborn, M., and Weber, K. (1986) *Trends Biochem. Sci.* **11**, 469–472
295. Gibson, M. A., Kumaratilake, J. S., and Cleary, E. G. (1989) *J. Biol. Chem.* **264**, 4590–4598
296. Steinert, P. M. (1991) *J. Structural Biol.* **107**, 157–174
297. Cohlberg, J. A. (1993) *Trends Biochem. Sci.* **18**, 360–362
298. Nakajima-Iijima, S., Hamada, H., Reddy, P., and Kakunaga, T. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6133–6137
299. Orlova, A., Chen, X., Rubenstein, P. A., and Egelman, E. H. (1997) *J. Mol. Biol.* **271**, 235–243
300. Fulton, A. B. (1984) *The Cytoskeleton: Cellular Architecture and Choreography*, Chapman and Hall, New York
301. Bershadsky, A. D., and Vasiliev, J. M. (1988) *Cytoskeleton*, Plenum, New York
302. DeRosier, D. J., Tilney, L. G., and Egelman, E. (1980) *Nature (London)* **287**, 291–296
303. Collier, N. C., and Wang, K. (1982) *J. Biol. Chem.* **257**, 6937–943
304. Moon, R. T., and Lazarides, E. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5494–5499
305. Meek, R. L., Lonsdale-Eccles, J. D., and Dale, B. A. (1983) *Biochemistry* **22**, 4867–4877
306. Steinert, P. M., and Idler, W. W. (1979) *Biochemistry* **18**, 5664–5669
307. Dustin, P. (1980) *Sci. Am.* **243**(Aug.), 67–76
308. Mandelkow, E., and Mandelkow, E.-M. (1994) *Current Opinion in Structural Biology* **4**, 171–179
309. Hyams, J. S., and Lloyd, C. W., eds. (1994) *Microtubules*, Wiley-Liss, New York
310. Sosa, H., and Milligan, R. A. (1996) *J. Mol. Biol.* **260**, 743–755
311. Margulis, L., To, L., and Chase, D. (1978) *Science* **200**, 1118–1124
312. Thormählen, M., Marx, A., Müller, S. A., Song, Y.-H., Mandelkow, E.-M., Aebi, U., and Mandelkow, E. (1998) *J. Mol. Biol.* **275**, 795–809
313. Weber, K., Pollack, R., and Bibring, T. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 459–463
314. Wolf, S. G., Nogales, E., Kikkawa, M., Gratzinger, D., Hirokawa, N., and Downing, K. H. (1996) *J. Mol. Biol.* **262**, 485–501
315. Nogales, E., Wolf, S. G., and Downing, K. H. (1998) *Nature (London)* **391**, 199–203
316. Pennisi, E. (1998) *Science* **279**, 176–177
317. Zheng, Y., Wong, M. L., Alberts, B., and Mitchison, T. (1995) *Nature (London)* **378**, 578–583
318. Hackney, D. D. (1995) *Nature (London)* **376**, 215–216

## References


319. Raff, E. C., Fackenthal, J. D., Hutchens, J. A., Hoyle, H. D., and Turner, F. R. (1997) *Science* **275**, 70–73
320. Angelastro, J. M., and Purich, D. L. (1992) *J. Biol. Chem.* **267**, 25685–25689
321. Terry, B. J., and Purich, D. L. (1980) *J. Biol. Chem.* **255**, 10532–10536
322. Rakic, P., Knyihar-Csillik, E., and Csillik, B. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9218–9222
323. Rodionov, V. I., and Borisy, G. G. (1997) *Science* **275**, 215–218
324. Hoenger, A., and Milligan, R. A. (1996) *J. Mol. Biol.* **263**, 114–119
325. Fan, J., Griffiths, A. D., Lockhart, A., Cross, R. A., and Amos, L. A. (1996) *J. Mol. Biol.* **259**, 325–330
326. Hirose, K., Fan, J., and Amos, L. A. (1995) *J. Mol. Biol.* **251**, 329–333
327. Knop, M., and Schiebel, E. (1997) *EMBO J.* **16**, 6985–6995
328. Moritz, M., Braunfeld, M. B., Sedat, J. W., Alberts, B., and Agard, D. A. (1995) *Nature (London)* **378**, 638–640
329. Berridge, M. J. (1990) *J. Biol. Chem.* **265**, 9583–9586
330. Arnold, C. S., Johnson, G. V. W., Cole, R. N., Dong, D. L.-Y., Lee, M., and Hart, G. W. (1996) *J. Biol. Chem.* **271**, 28741–28744
331. Bosc, C., Cronk, J. D., Pirollet, F., Watterson, D. M., Haiech, J., Job, D., and Margolis, R. L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2125–2130


## Study Questions

1. Rewrite Equations 6-75 through 6-77 in terms of dissociation constants. These may be labeled $K_1$, $K_2$, $K_i$, etc., as is conventional, but you may prefer to use $K_{1d}$, $K_{2d}$, $K_{id}$, etc., to avoid confusion.

2. A molecule has two identical binding sites for a ligand X. The Gibbs energy of interaction between ligands bound to the same molecule, ε, is defined as the change in Gibbs energy of binding of the ligand to the molecule that results from the prior binding of a ligand at the adjacent site. If the saturation fraction is Y, show from the equation for the binding isotherm that the following equation holds when $Y = 1/2$:

$$dY/d\ln[X] = \frac{1}{2(1 + e^{\varepsilon/RT})}$$

3. The hydrogen ion binding curve for succinate is shown is Fig. 7-4. From the curve estimate ε and the microscopic association constants.

4. A linear chain molecule has a very large number of identical binding sites for a ligand X. The Gibbs energy of interaction between ligands bound to adjacent sites is ε. Interactions between non-nearest neighbors are considered negligible. If the binding constant for a site adjacent to unoccupied sites is $K_r$, the binding isotherm is given by

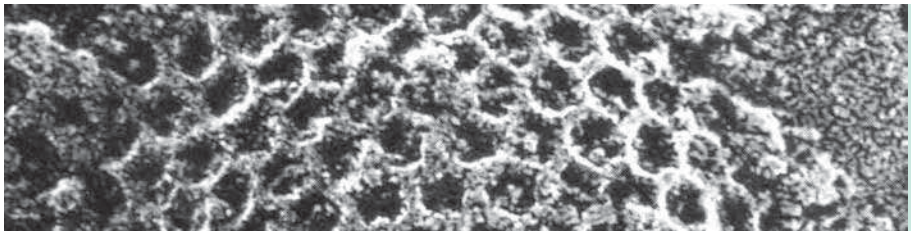
$$Y = \frac{1}{2} + \frac{K[X]e^{-\varepsilon/RT} - 1}{2\{K[X]e^{-\varepsilon/RT} - 1)^2 + 4K[X]\}^{1/2}}$$

   [Applequist, J. (1977) *J. Chem. Ed.* **54**, 417]. Show from the equation for the binding isotherm that the following equation holds at $Y = 1/2$;

$$dY/d\ ln\ [X] = 1/4e^{\varepsilon/2RT}$$

5. The binding of adenosine to polyribouridylic acid [poly(U)] has been studied by the method of equilibrium dialysis [Huang and Ts'o (1966) *J. Mol. Biol.* **16**, 523]. The table below gives the fraction of poly(U) sites occupied, Y at various molar concentrations of free adenosine [A] at 5°C. Assuming that the nearest-neighbor interaction model is correct, determine the intrinsic association constant for the binding of adenosine to poly U and the free energy of interaction of adjacent bound adenosines. Do the bound molecules attract or repel each other?

| [A] x $10^3$ | Y | [A] x $10^3$ | Y |
|---|---|---|---|
| 0.51 | 0 | 3.07 | 0.72 |
| 2.10 | 0 | 4.00 | 0.92 |
| 2.70 | 0.15 | 6.50 | 0.93 |
| 2.96 | 0.36 | 8.50 | 0.93 |
| 3.01 | 0.52 | 10.00 | 1.00 |



Portion of an endocytic vacuole forming in the plasma membrane of a cultured fibroblast. The view is from the inside of the cell and shows a large clathrin cage assembling to form a coated vesicle. The overall diameter of the vacuole is ~0.2 µm. Clathrin cages vary in size and in the number of faces but are typically ~0.1 µm in diameter (see Fig. 8-27 and associated references). Courtesy of Barbara Pearse.

# Contents


380 **A. Lipid Structures**
380 1. Fatty Acids, Fatty Alcohols, and Hydrocarbons
382 2. Acylglycerols, Ether Lipids, and Waxes
382 3. Phospholipids
387 4. Glycolipids
389 5. Sphingolipids
389 6. Sterols and Other Isoprenoid Lipids
390 **B. Membranes**
390 1. The Structure of Membranes
392 *Lipids of membranes*
392 *Liquid crystals, liposomes, and artificial membranes*
394 *Physical properties of membrane lipids*
396 *Functions of phospholipid head groups*
397 *Non-bilayer structures of phospholipids*
397 *Membrane fluidity and life*
400 *Electrical properties of membranes*
401 *The two sides of a membrane*
401 2. Membrane Proteins
401 *Integral membrane proteins*
402 *Anchors for proteins*
403 *Analyzing erythocyte membranes*
404 *Glycoproteins*
405 *Connections to the cytoskeleton*
405 *Integrins and focal adhesions*
406 *Other cell adhesion molecules*
407 *Peripheral proteins of the outer membrane surface*
409 3. Enzymes and Membrane Metabolism
410 **C. The Transport of Molecules through Membranes**
410 1. Facilitated Diffusion and Active Transport
411 2. Pores, Channels, and Carriers
411 *Porins*
411 *Aquaporins*
412 *Ion channels*
414 *Channel-forming toxins and antibiotics*
414 *Ionophores and other mobile carriers*
415 3. The 12-Helix Major Facilitator Superfamily
415 *Entrance of sugars into cells*
416 *Cotransport of sugars and other nutrients with $H^+$ or $Na^+$*
417 4. Active Transport Systems
417 *Periplasmic permeases*
419 *The bacterial phosphotransferase system*
420 5. Transport of Ions
420 *Anions*
421 *Cation channels*
422 *Active transport of cations*
422 *ATP-driven ion pumps*
425 6. Exocytosis, Endocytosis, and the Flow of Membrane Constituents
427 **D. Communication**
427 **E. The Extracellular Matrix and Cell Walls**
427 1. The Structure of Bacterial Cell Walls
428 *The outer membrane of gram-negative bacteria*
431 *Teichoic and teichuronic acids*
431 *Mycobacteria*
431 *Other bacterial coats*
431 2. The Surroundings of Animal Cells
432 *The collagens*
436 *Elastic fibers*
436 *Cartilage and basement membranes*
440 *Fibrillin and Marfan's syndrome*
440 *The cuticles of invertebrates*
440 *Bones, teeth, and shells*
443 3. Cell Walls of Fungi and Green Plants

444 **References**
452 **Study Questions**

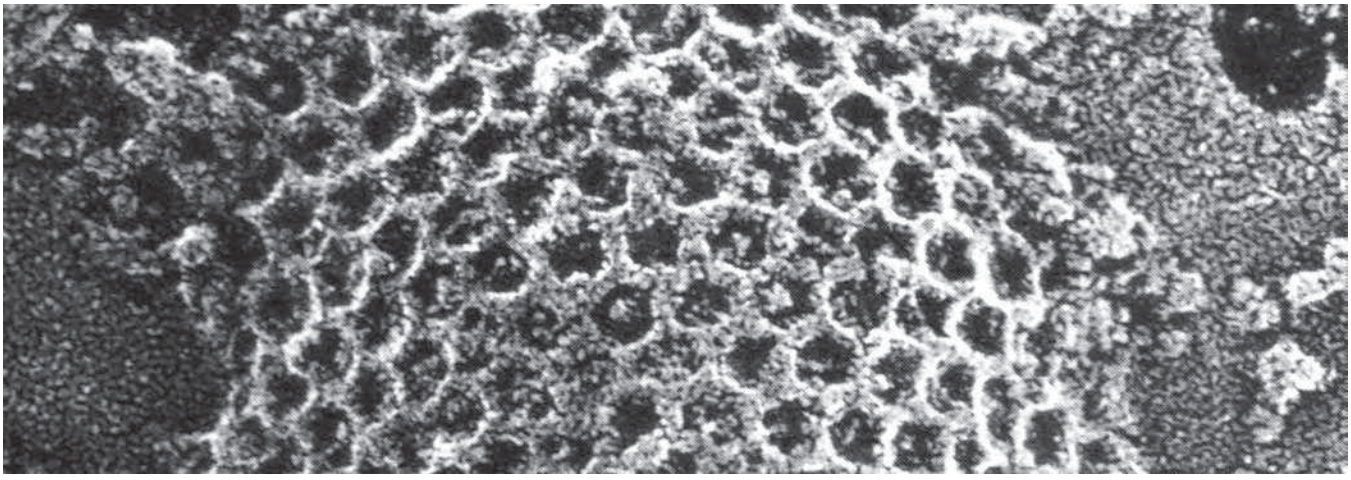
**Boxes**

384 Box 8-A The Platelet-Activating Factor
386 Box 8-B Dipalmitoylphosphatidylcholine and the Surfactant System of the Lungs
398 Box 8-C Electron Paramagnetic Resonance (EPR) Spectra and "Spin Lables"
418 Box 8-D Colicins: Antibiotic Proteins
438 Box 8-E Genetic Defects of Collagen Structure
439 Box 8-F Skin
442 Box 8-G The Biochemistry of Teeth

**Tables**

380 Table 8-1 Some Important Fatty Acids
381 Table 8-2 Fatty Acid Composition (in %) of Some Typical Fats and Oils
393 Table 8-3 Estimated Chemical Compositions of Some Membranes
435 Table 8-4 Types of Vertebrate Collagen

# 8 Lipids, Membranes, and Cell Coats



The boundary between a living cell and its surroundings is the incredibly thin (7–10 nm) plasma membrane. This vital partition, which controls the flow of materials into and out of a cell and which senses and controls the response of cells to hormones and other external signals, consists largely of **phospholipids** together with embedded proteins. The nonpolar chains of the phospholipids stick together to form a double molecular layer or **bilayer** which provides the basic structure of almost all biological membranes.

Phospholipids, together with other natural materials that have a high solubility in apolar solvents or are structurally related to compounds with such solubility properties, are classified as **lipids**. The most abundant lipids are the fats, compounds that are stored by animals and by many plants as an energy reserve (Fig. 8-1). Other lipids form the outer cuticle of plants and yet others serve as protective coatings on feathers and hair. Vitamins A, D, K, and E and ubiquinone are all lipids as are a variety of hormones and such light-absorbing plant pigments as the chlorophylls and carotenoids. Many of these compounds are dissolved in or partially embedded in the plasma membrane of bacteria or in the mitochondrial and chloroplast membranes of higher organisms. Membranes serve many purposes. The most obvious is to divide space into compartments. Thus, the plasma membrane forms cell boundaries and mitochondrial membranes separate the enzymes and metabolites of mitochondria from those of the cytosol. Membranes are semipermeable and regulate the penetration into cells and organelles of both ionic and nonionic substances. Many of these materials are brought into the cell against a concentration gradient. Hence, osmotic work must be done in a process known as **active transport**. Many enzymes,

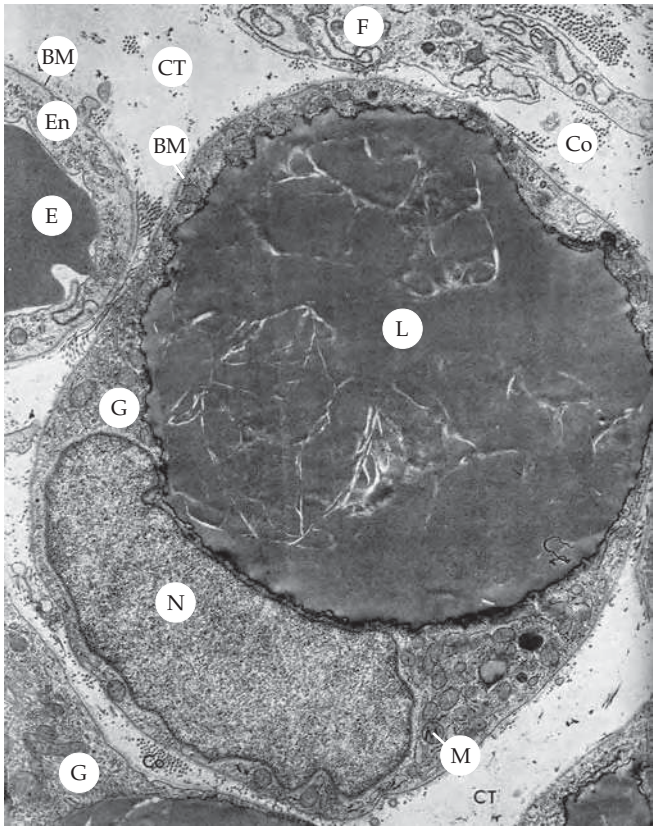


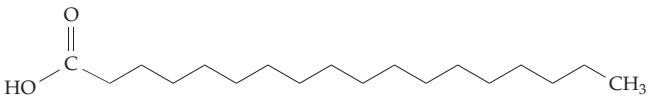

**Figure 8-1** Electron micrograph of a thin section of a fat storage cell or adipocyte. L, the single large fat droplet; N, nucleus; M, mitochondria; En, endothelium of a capillary containing an erythrocyte (E); CT, connective tissue ground substance which contains collagen fibers (Co) and fibroblasts (F). The basement membranes (BM) surrounding the endothelium and the fat cell are also marked. From Porter and Bonneville.[6] Courtesy of Mary Bonneville.

including those responsible for most of the oxidative metabolism of cells, are found in membranes of bacteria and of mitochondria. Within the chloroplasts of green leaves, highly folded membranes containing chlorophyll absorb energy from the sunlight. Thin membranes contain the photoreceptor proteins that function in vision. Electrical impulses are transmitted along the membranes of nerve cells.

The outer surfaces of membranes are designed to interact with the cell's external world. Special receptors sense the presence of hormones. Binding proteins await the arrival of needed nutrients and help to bring them into cells. Highly individual arrangements of protein and of the carbohydrate "fuzz" of glycoproteins and sphingolipids screen the outer surface, helping to prevent attack by foreign bacteria, viruses, and toxins.

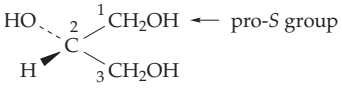
## A. Lipid Structures

Unlike proteins, polysaccharides, and nucleic acids, most lipids are not polymers. However, they are made by linking together smaller molecules.[1–5] Among the "building blocks" of lipids are **fatty acids, glycerol, phosphoric acid**, and **sugars**. Many lipids have both polar and nonpolar regions. This gives them an **amphipathic** character, i.e., a tendency toward both hydrophobic and hydrophilic behavior, and accounts for their tendency to aggregate into membranous structures. Notice that in the following structure the carbon atoms of glycerol have been numbered 1 to 3. Although glycerol is *not* chiral, the positioning of the two hydroxymethyl groups is not equivalent. If the *priority*, used in the *RS* system (page 42) for the C-1 group is raised (e.g., by ester formation) to be higher than that for the C-3 group, the molecule would have the *S* configuration. The C-1 group is said to be pro-*S* and the C-3 group pro-*R*. According to the stereochemical numbering (*sn*) system, which is discussed further in Chapter 9, the carbon in the pro-*S* position is numbered 1.




A fatty acid

Stearic acid, 18 carbon atoms, major component of animal triacylglycerols




Glycerol

Labeled by the stereochemical numbering (*sn*) system

## 1. Fatty Acids, Fatty Alcohols, and Hydrocarbons

The names and structures of some fatty acids are summarized in Table 8-1. Notice that these acids have straight carbon chains and may contain one or more double bonds. Except for the smallest members of the series, which are soluble in water, fatty acids are strongly hydrophobic. However, they are all acids with p$K_a$ values in water of ~4.8. To the extent that free fatty acids occur in nature, they are likely to be found in interfaces between lipid and water with the carboxyl groups dissociated and protruding into the water. However, most naturally occurring fatty acids

### TABLE 8-1
### Some Important Fatty Acids

| No. carbon atoms | Systematic name | Common name | Abbreviation[a] | Common name of acyl group |
|---|---|---|---|---|
| | | **Saturated fatty acids** | | |
| 1 | Methanoic | Formic | | Formyl |
| 2 | Ethanoic | Acetic | | Acetyl |
| 3 | Propanoic | Propionic | | Propionyl |
| 4 | Butanoic | Butyric | 4:0 | Butyryl |
| 12 | Dodecanoic | Lauric | 12:0 | Lauroyl[b] |
| 14 | Tetradecanoic | Myristic | 14:0 | Myristoyl |
| 16 | Hexadecanoic | Palmitic | 16:0 | Palmitoyl |
| 18 | Octadecanoic | Stearic | 18:0 | Stearoyl |
| 20 | Eicosanoic | Arachidic | 20:0 | |
| 22 | Docosanoic | Behenic | 22:0 | |
| 24 | Tetracosanoic | Lignoceric | 24:0 | |
| | | **Unsaturated fatty acids[c]** | | |
| 4 | | Crotonic | 4:1(2*t*) | Crotonoyl |
| 16 | | Palmitoleic | 16:1(9*c*) | |
| 18 | | Oleic | 18:1(9*c*) | Oleoyl |
| 18 | | Vaccenic | 18:1(11*c*) | |
| 18 | | Linoleic | 18:2(9c,12*c*) | |
| 18 | | Conjugated linoleic | 18:2(9c,11*t*) | |
| 18 | | Linolenic | 18:3(9*c*,12*c*,15*c*) | |
| 20 | | Arachidonic | 20:4(5*c*,8*c*,11*c*,14*c*) | |

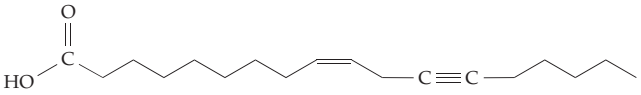
[a] The number of carbon atoms is given first, then the number of double bonds. The positions of the lowest numbered carbon of each double bond and whether the configuration is *cis* (*c*) or *trans* (*t*) are indicated in parentheses.

[b] Official IUPAC names of these and other acyl groups have been designated by the Commission of the Nomenclature of Organic Chemistry in *Pure and Applied Chemistry* **10**, 111–125 (1965). In a number of cases IUPAC inserted an *o* in the traditional name, e.g., palmityl became palmitoyl and crotonyl became crotonoyl. However, acetyl was not changed. In many cases the systematic names, e.g., hexadecanoyl (from hexadecanoic acid), are preferable and IUPAC–IUB recommends that alkyl radicals always be designated by systematic names, e.g., hexadecyl, *not* palmityl alcohol. The older use of palmityl for both acyl and alkyl radicals was one reason for IUPAC's adoption of new names for acyl radicals.

[c] Systematic names are not often used because of their complexity, e.g., linolenic acid is *cis*,*cis*,*cis*-9,12,15-octadecatrienoic acid.

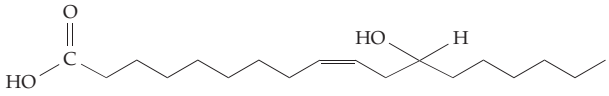
are esterified or combined via amide linkages in complex lipids. For example, ordinary fats are largely the fatty acid esters of glycerol called **triacylglycerols** (triglycerides).

There is a seemingly endless variety of fatty acids, but only a few of them predominate in any single organism. Most fatty acid chains contain an even number of carbon atoms. In higher plants the $C_{16}$ **palmitic acid** and the $C_{18}$ unsaturated **oleic** and **linoleic acids** predominate. The $C_{18}$ saturated **stearic acid** is almost absent from plants and $C_{20}$ to $C_{24}$ acids are rarely present except in the outer cuticle of leaves. Certain plants contain unusual fatty acids which may be characteristic of a taxonomic group. For example, the Compositae (daisy family) contain acetylenic fatty acids and the castor bean contains the hydroxy fatty acid **ricinoleic acid**.



Crepenynic acid 18:2 (9c, 12a)

Accounts for 60% of the fatty acids in seeds of *Crepis foetida*, a member of the compositae family



Ricinoleic acid (12-hydroxyoleic acid)

Accounts for up to 90% of the fatty acids of *Ricinus communis* (castor bean)

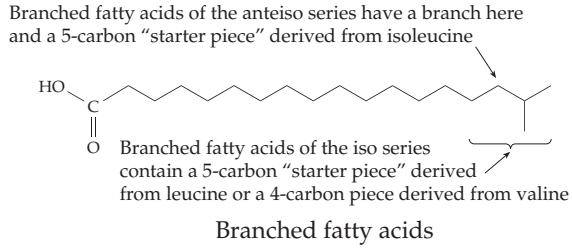
Like plants, animals contain palmitic and oleic acids. In addition, large amounts of stearic acid and small amounts of the $C_{20}$, $C_{22}$, and $C_{24}$ acids are also present. Phospholipids of photoreceptor membranes of the retina contain fatty acid chains as long as $C_{36}$.[7] The variety of fatty acids found in animals is greater than in a given plant species. A large fraction of the fatty acids present in most higher organisms are *unsaturated* and contain strictly *cis* double bonds. Table 8-2 shows the fatty acid composition of some typical triacylglycerol mixtures.

**TABLE 8-2**
**Fatty Acid Composition (in %) of Some Typical Fats and Oils**[a]

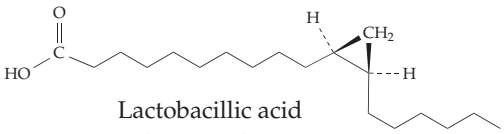
| | No. of carbon atoms and (following colon) the number of double bonds | | | | | |
|---|---|---|---|---|---|---|
| Fats and oils | 14 | 16 | 18 | 16:1 | 18:1 | 18:2 |
| Human depot fat | 3 | 23 | 4 | 8 | 45 | 10 |
| Beef tallow | 4 | 30 | 25 | 5 | 36 | 1 |
| Corn oil | | 13 | 2 | | 31 | 54 |
| Lard | 1 | 28 | 15 | 3 | 42 | 9 |

[a] From Gunstone.[8]

Bacteria usually lack polyunsaturated fatty acids but often contain branched fatty acids, cyclopropane-containing acids, hydroxy fatty acids, and unesterified fatty acids. Mycobacteria, including the human pathogen *Mycobacterium tuberculosis*, contain **mycolic acids**. In these compounds the complex grouping R contains a variety of functional groups including –OH, $-OCH_3$, C=O, –COOH, cyclopropane rings, methyl branches, and C=C bonds. Each species of *Mycobacterium* contains about two dozen different mycolic acids[9,10] as well as other complex $C_{30}$–$C_{56}$ fatty acids (see Eq. 21-5).[11]

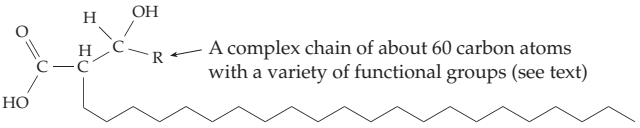



Branched fatty acids



Lactobacillic acid

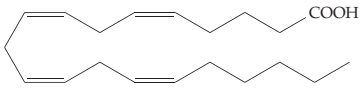
A major fatty acid of lactobacilli




$C_{22}$ or $C_{24}$ alkyl group

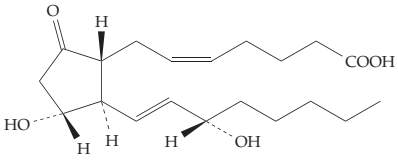
A mycolic acid

Certain polyunsaturated fatty acids are essential in the human diet (see Box 21-B). One of these, **arachidonic acid** (which may be formed from dietary linoleic acid), serves as a precursor for the formation of the hormones known as **prostaglandins** and a series of related **prostanoids**. Lipids of animal origin also



Arachidonic acid

A nutritionally essential fatty acid (Box 21-A) is shown here in a folded conformation

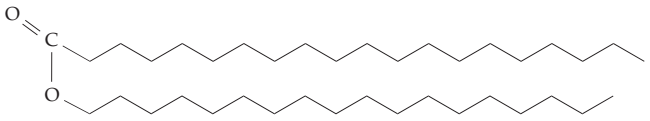


Prostaglandin $PGE_2$

One of a family of hormones, this one is synthesized in tissues from arachidonic acid

contain unusual unsaturated fatty acids. Among them, **conjugated linoleic acids** are receiving attention for their possible cancer-preventive action. The predominant form in meats, dairy products, and the human body is the $C_{18}$ 9-*cis*, 11-*trans* isomer whose two double bonds are conjugated.[11a]

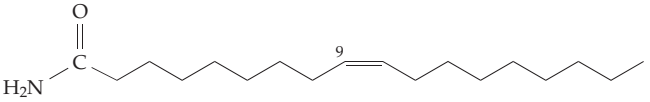
Other lipid components include the **fatty alcohols** which are formed by reduction of the acids. These are esterified with fatty acids to form **waxes**. Both fatty alcohols and free fatty acids occur in waxes together with the esterified forms. These mixtures are found on exterior surfaces of plants and animals. Plants and, to a limited extent, animals are able to decarboxylate fatty acids in a multistep process to **alkanes** and these too are important constituents of some waxes. Small amounts of fatty acid amides such as *cis*-9,10-octadecenoamide are present in low concentrations in



A wax

An ester of a fatty acid and a fatty alcohol

the cerebrospinal fluid of cats and rats as well as humans. This compound accumulates in cats that are deprived of sleep. When the synthetic compound was injected into rats they fell into apparently normal sleep.[12]

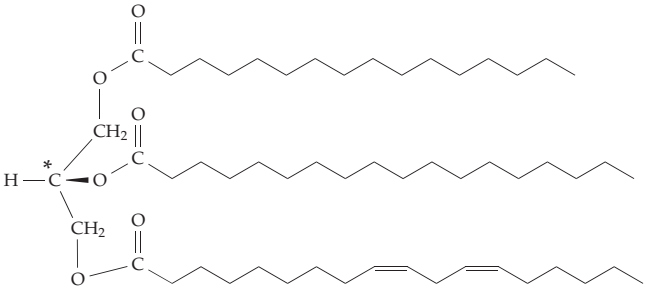


*cis*-9,10-Octadecenoamide (Oleic acid amide)

Insects make unsaturated as well as saturated hydrocarbons. The former as well as long-chain alcohols and their esters often form the volatile **pheromones** with which insects communicate. Thus, the female pink bollworm attracts a male with a sex pheromone consisting of a mixture of the *cis*,*cis* and *cis*,*trans* isomers of 7,11-hexadecadienyl acetate,[13] and European corn borer males are attracted across the cornfields of Iowa by *cis*-11-tetradecenyl acetate.[14] Addition of a little of the *trans* isomer makes the latter sex attractant much more powerful. Since more than one species uses the same attractant, it is possible that the males can distinguish between different ratios of isomers or of mixtures of closely related substances.

## 2. Acylglycerols, Ether Lipids, and Waxes

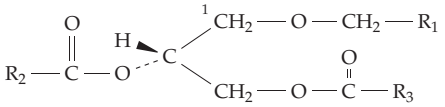
The components of complex lipids are linked in a variety of ways. Often, glycerol acts as the central unit, e.g., combining in ester linkage with three fatty acids to form **triacylglycerols** (triglycerides), the common fats of adipose tissues and plant oils. Diacyl- and monoacylglycerols (diglycerides, and monoglycerides) are



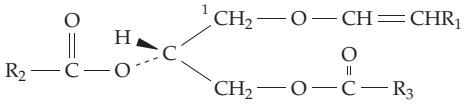
A triacylglycerol (fat)

Notice the chiral center, designated by the asterisk

present to a lesser extent. In addition, small amounts of **alkyl ethers** or **alkenyl ethers** are often present in isolated lipids. They are especially abundant in fish liver oils.



1-Alkyl-2,3-diacyl-*sn*-glycerol, an alkyl ether lipid



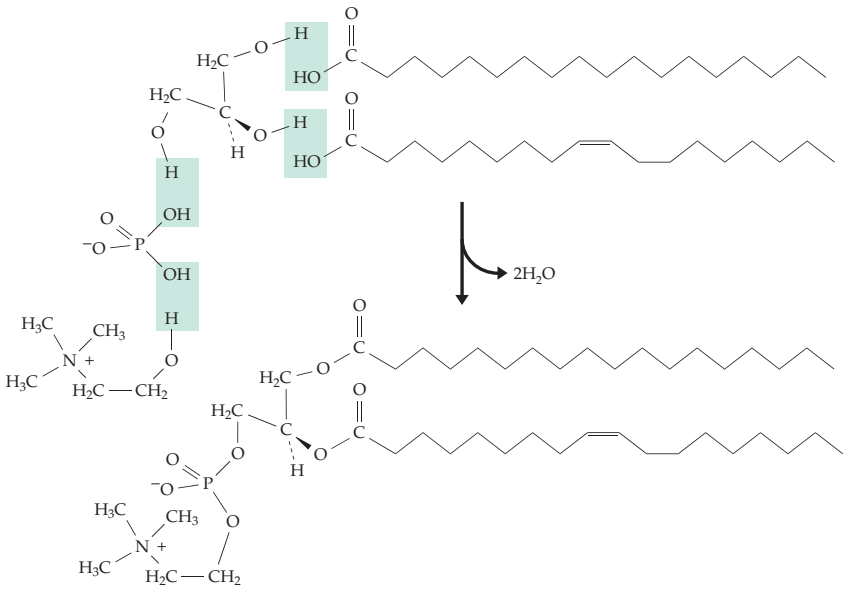
1-Alkenyl-2,3-diacyl-*sn*-glycerol, an alkenyl ether

These ether lipids are all chiral molecules with an *R* configuration but are derivatives of the nonchiral glycerol. The carbon atoms of glycerol are numbered using the stereochemical system which is described on p. 470. The ether linkage is to the *sn*-1 carbon atom. Most phospholipids are derivatives of the *sn*-3 phosphate ester of glycerol.

Triacylglycerols and the ether lipids described in the previous section are classified as **neutral lipids**. Other neutral lipids are alcohols, waxes, aldehydes, and hydrocarbons derived from fatty acids. These sometimes have specific biological functions. For example, fatty aldehydes are important in the bioluminescence of bacteria (Eq. 23-47).

## 3. Phospholipids

As major constituents of biological membranes, phospholipids play a key role in all living cells. The two principal groups of phospholipids are the




A phosphatidylcholine (lecithin). (8-1)

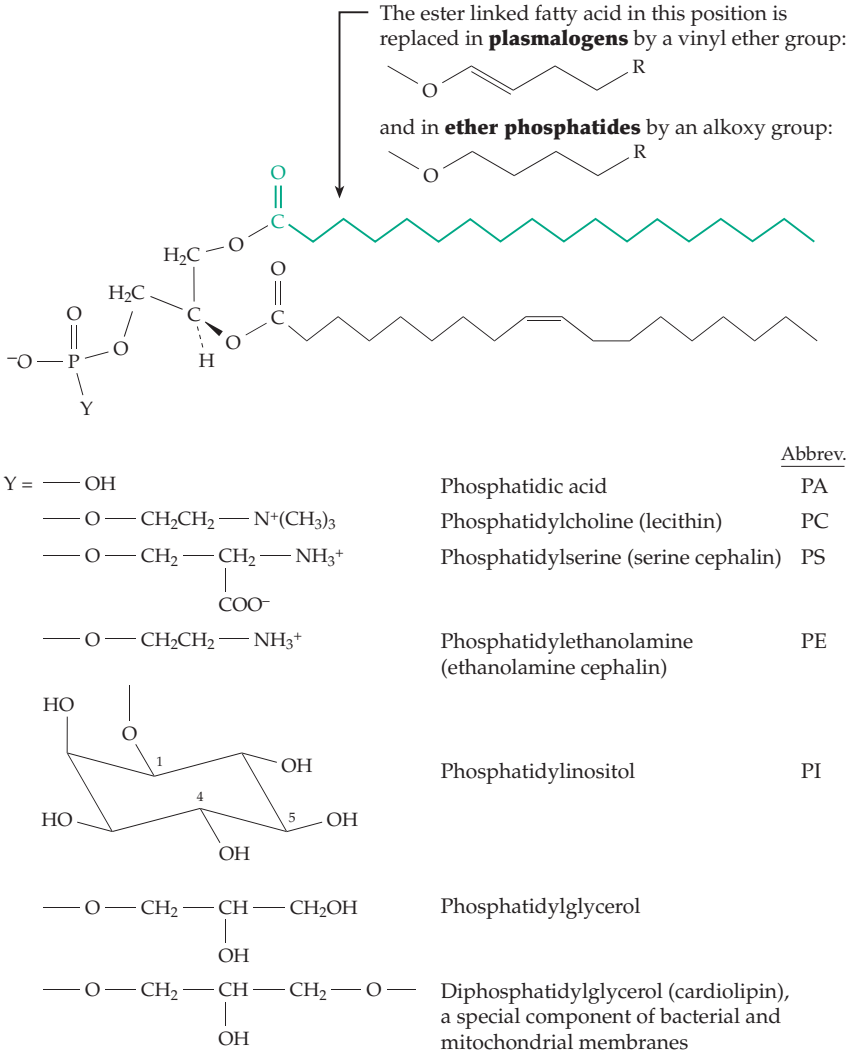
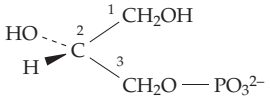



**Figure 8-2** Structures of some phosphatides (glycerophospholipids).

**glycerophospholipids** (glycerophosphatides; Fig. 8-2) which contain the alcohol glycerol and the **sphingophospholipids** which contain the alcohol **sphingosine** (Section 5). The glycerophospholipids can be thought of as arising from the building blocks glycerol, fatty acids, the dihydrogen phosphate ion $H_2PO_4^-$, and the appropriate alcohol by removal of four molecules of water (Eq. 8-1). They are derivatives of *sn*-glycerol-3-phosphate. Esterification of this alcohol with two fatty acids gives a **phosphatidic acid** (Fig. 8-2). Formation of a phosphate diester linkage to one of the alcohols



*sn*-Glycerol 3-phosphate (L-α-glycerol phosphate)

choline, serine, or ethanolamine yields a glycerophospholipid. The resulting three groups of phospholipids are called **phosphatidylcholine** (lecithin), **phosphatidylserine**, and **phosphatidylethanolamine**, respectively (Fig. 8-2).

The phosphate and choline, ethanolamine, or serine portions of the phosphatide are electrically charged and provide a polar "head" for the molecule. In all three cases the positively charged group is able to fold back and form an ion pair with the negatively charged phosphate group. However, the methyl groups surrounding the nitrogen in phosphatidylcholine prevent a very close approach and with phosphatidylserine the adjacent carboxylate group weakens this electrostatic interaction. Unlike the triacylglycerols, most of which are liquid at body temperature, phospholipids are solid at this temperature. This property, like the ionic properties of the phosphatides, is doubtless related to their suitability for functioning in biological membranes.

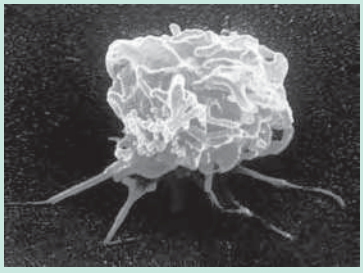
Lecithins and related phospholipids usually contain a saturated fatty acid in the C-1 position but an unsaturated acid, which may contain from one to four double bonds, at C-2. Arachidonic acid is often present here. Hydrolysis of the ester linkage at C-2 yields a 1-acyl-3-phosphoglycerol, better known as a **lysophosphatidylcholine**. The name comes from the powerful detergent action of these substances which leads to lysis of cells. Some snake venoms contain phospholipases that form lysophosphatidylcholine. Lysophosphatidic acid (1-acyl-glycerol-3-phosphate) is both an intermediate in phospholipid biosynthesis (Chapter 21) and also a signaling molecule released into the bloodstream by activated platelets.[15]

Another group of phosphatides contain the hexahydroxycyclohexane known as **inositol** (Fig. 8-2, see also Chapter 21).[16] Phosphatidylinositol, as well as smaller amounts of phosphatides derived from phosphate esters of inositol are present in membranes of all eukaryotes and have a specific role in regulating responses of cells to hormones and other external agents. See Chapter 11 for details. Phosphatidylinositol also forms part of "anchors" used to hold certain proteins onto membrane surfaces (see Fig. 8-13).

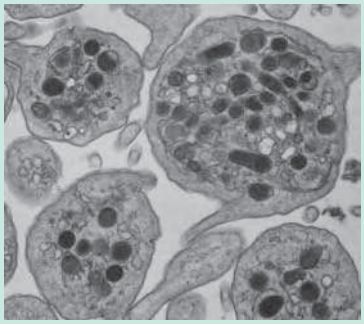
Bacteria and plants often make the anionic **phosphatidyglycerol** in which the second glycerol is esterified at its *sn*-1 position with the phosphate. Bacteria, as well as mitochondria, contain diphosphatidylglycerol (**cardiolipin**) in which phosphatidyl groups are attached at both the 1 and 3 positions of glycerol (Fig. 8-2). Ether phospholipids, analogous to the ether lipids described in Section 2, are also widely distributed. The alkenyl ether analogs of phosphatidylcholine (Fig. 8-2) are called **plasmalogens**.[17] In neutrophils the 1-*O*-alkyl ethers contain the major share of the cell's arachidonic acid, which is esterified in the 2 position.[18,19]

In halophilic (salt loving), thermophilic, and methanogenic bacteria, most of the lipids present are either

## BOX 8-A THE PLATELET-ACTIVATING FACTOR



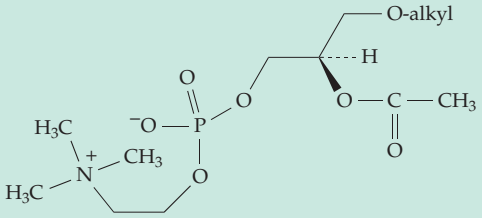
Scanning EM of activated blood platelets. © Quest, Photo Researchers



Transmission EM of thin section through activated blood platelets. © Photo Researchers

Part of the immune response consists of the release from stimulated neutrophils, macrophages, and other cells of a **platelet activating factor (PAF)**, a material that "activates" blood platelets. Activated platelets aggregate, a response that initiates clot formation. They may also be lysed and release stored substances that include platelet-derived growth factors (see Chapter 11) and fibrin stabilizing factor, a proenzyme that is converted to the protein-crosslinking enzyme transglutaminase (Eq. 2-17). See Fig. 12-17. However, the principal interest in the platelet-activating factor has arisen from its powerful effect in inducing **inflammation** in surrounding tissues.

PAF has been identified as the following simple ether phospholipid:[a–d]

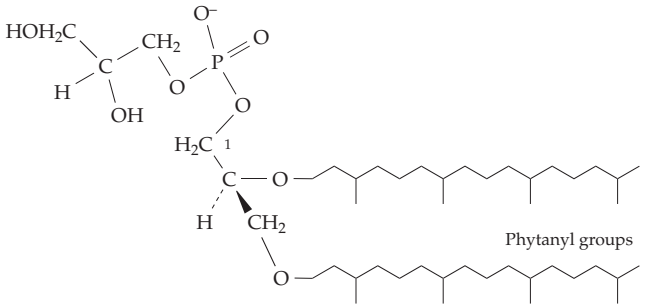



Platelet-activating factor (PAF)
1-*O*-alkyl-2-acetyl-*sn*-glycerophosphocholine

A remarkably potent compound, its effects on platelets are observed at concentrations of $10^{-11}$ to $10^{-10}$ M. Both lyso-PAF and the glycero-1- phosphocholine enantiomer are inactive. Specific receptors for this factor are evidently present on platelet surfaces.[e–g]

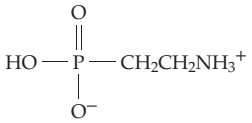
One effect of PAF on platelets is to induce a rapid (5–10s) cleavage of phosphatidylinositol 4,5-bisphosphate by phospholipase C to give **diacylglycerol** and **inositol 1,4,5-trisphosphate**. The subsequent effects of these two substances in causing a rapid influx of $Ca^{2+}$ and in inducing a series of secondary responses are outlined in Fig. 11-9. Among these responses are the release of the materials stored in the platelet's granules. PAF also appears to inhibit adenylate cyclase[b] and causes vasodilation, a property not expected for a compound that stimulates clotting. Receptors for PAF are also present in the brain, where this phosholipid may function in regulation of development.[h]

phospholipids and glycolipids containing the $C_{20}$ isoprenoid **phytanyl** group or the $C_{40}$ **diphytanyl** group[20–25] (see also Section 4), related isoprenoid alcohols, or long-chain 1,2-diols.[25] An example of a diphytanylglycerophospholipid is the following:




A phospholipid derived from 2,3-di-*O*-phytanyl-*sn*-glycerol

Many other phospholipids are present in small amounts or in a limited number of species. These include **phosphonolipids**, which contain a C–P bond and are abundant in ciliate protozoa such as *Tetrahymena* and in some other invertebrates.[26] Phosphonoethylamine replaces phosphoethanolamine in these lipids. A consequence of this structural alteration is a high



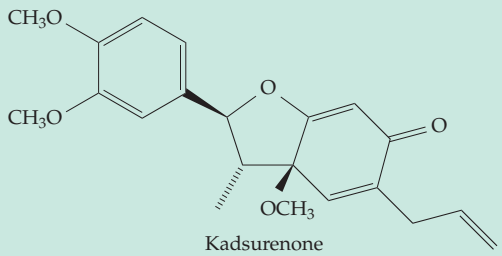

Phosphonoethylamine

degree of resistance to the action of the enzyme phospholipase C. The phosphonolipids of the external membrane of *Tetrahymena* are also ether lipids with an alkoxy group in the *sn*-1 position. This makes them

## BOX 8-A (continued)

Stimulated platelets release arachidonic acid rapidly from their phospholipids, apparently as a result of activation of phospholipase $A_2$. The released arachidonate can in turn be metabolized to endoperoxides and thromboxane $A_2$ (Chapter 21). These compounds are also potent activators of platelets and cause a self-activating or **autocrine** effect.[i,j] While PAF has a beneficial function, it can under some conditions contribute in a dangerous way to inflammation and to allergic responses including **anaphylaxis**,[j] **asthma**[g] and **cold-induced urticaria**.[k] Although the effect of PAF is separate from those of histamine and of leukotrienes, these agents may act cooperatively to induce inflammation.[l]

The biosynthesis of PAF is discussed in Chapter 21, Section C. One pathway is deacylation at the glycerol C-2 of a longer chain alkyl phospholipid followed by acetylation at the same position.[m,n] Platelets can inactivate PAF by the inverse sequence: deacetylation followed by acylation with arachidonic acid. PAF can also be hydrolyzed by a PAF acetylhydrolase, a phospholipase.[o] Absence of this enzyme in the brain may be related to a human brain malformation.[h] Some tissues may contain a lipid that inhibits binding of PAF.[p] The following compound from a Chinese herb binds to PAF receptors and may be the forerunner of useful drugs.[q]




Kadsurenone


[a] Cusack, N. J. (1980) *Nature (London)* **285**, 193
[b] Hanahan, D. J. (1986) *Ann. Rev. Biochem.* **55**, 483–509
[c] Prescott, S. M., Zimmerman, G. A., and McIntyre, T. M. (1990) *J. Biol. Chem.* **265**, 17381–17384
[d] Winslow, C. M., and Lee, M. L., eds. (1987) *New Horizons in Platelet Activating Factor Research*, Wiley, New York
[e] Hwang, S.-B., Lam, M.-H., and Pong, S.-S. (1986) *J. Biol. Chem.* **261**, 532–537
[f] Ishii, I., Izumi, T., Tsukamoto, H., Umeyama, H., Ui, M., and Shimizu, T. (1997) *J. Biol. Chem.* **272**, 7846–7854
[g] Bazan, N. G. (1995) *Nature (London)* **374**, 501–502
[h] Hattori, M., Adachi, H., Aoki, J., Tsujimoto, M., Arai, H., and Inoue, K. (1995) *J. Biol. Chem.* **270**, 31345–31352
[i] Bussolino, F., Sironi, M., Bocchietto, E., and Mantovani, A. (1992) *J. Biol. Chem.* **267**, 14598–14603
[j] Darius, H., Lefer, D. J., Smith, J. B., and Leefer, A. M. (1986) *Science* **232**, 58–60
[k] Grandel, K. E., Farr, R. S., Wanderer, A. A., Eisenstadt, T. C., and Wasserman, S. I. (1985) *N. Engl. J. Med.* **313**, 405–409
[l] Tomeo, A. C., Egan, R. W., and Durán, W. N. (1991) *FASEB J.* **5**, 2850–2855
[m] Billah, M. M., Eckel, S., Myers, R. F., and Siegel, M. I. (1986) *J. Biol. Chem.* **261**, 5824–5831
[n] Lee, T.-c, Ou, M.-c, Shinozaki, K., Malone, B., and Snyder, F. (1996) *J. Biol. Chem.* **271**, 209–217
[o] Stafforini, D. M., McIntyre, T. M., Zimmerman, G. A., and Prescott, S. M. (1997) *J. Biol. Chem.* **272**, 17895–17898
[p] Miwa, M., Hill, C., Kumar, R., Sugatan, J., Olson, M. S., and Hanahan, D. J. (1987) *J. Biol. Chem.* **262**, 527–530
[q] Hwang, S.-B., Lam, M.-H., Biftu, T., Beattie, T. R., and Shen, T.-Y. (1985) *J. Biol. Chem.* **260**, 15639–15645

## BOX 8-B DIPALMITOYLPHOSPHATIDYLCHOLINE AND THE SURFACTANT SYSTEM OF THE LUNGS

When air is exhaled the small alveoli of the lungs could collapse if it were not for the surface active material (surfactant) present in the fluid that coats the lungs.[a–e] In fact, the lack of adequate surfactant is the cause of **respiratory distress syndrome**, a major cause of death among premature infants and a disease that may develop in acute form in adults. The surfactant material forms a thin film of high fluidity at the air–liquid interface and lowers the surface tension from the 72 mN/m of pure water to <10 mN/m.[f,g] (Pay attention to the definition of surface tension.[h]) About 65% by weight of the surfactant is lecithin, mostly dipalmitoylphosphatidylcholine (see Fig. 8-4), a phospholipid resistant to attack by oxygen. Phosphatidylglycerol, in an unusually high concentration, accounts for ~ 12% of the human surfactant. Other phospholipids, plasmalogen,[i] cholesterol, proteins, and calcium ions are also present.

The surfactant contains four unique proteins, designated SP-A, SP-B, SP-C, and SP-D.[c,e] The major protein (SP-A) is a sialic acid-rich glycoprotein derived from a 26-kDa peptide, which contains a short collagen-like domain.[e,j,k] Like collagen, this domain contains glycosylated hydroxyproline. The C-terminal domain is a $Ca^{2+}$-dependent C-type lectin (Chapter 4), while the N-terminal domain is involved in oligomer formation through disulfide bridges. The overall structure is similar to that of the complement protein C1q (Chapter 31).[e,j] Protein D is also collagen-like[l] but evidently plays a very different functional role than SP-A. The latter associates with the major surfactant lipids but SP-D does not. It does bind phosphatidylinositol[m] and glucosylceramide, lipids that are present in small amounts. Perhaps SP-D helps to remove these polar lipids which might interfere with surfactant action.[e] Both proteins A and D may also have functions in the immune system.[l]

Proteins SP-B and SP-C are small extremely hydrophobic polypeptides consisting of 79 and 35 amino acid residues, respectively.[n,o] Aliphatic branched amino acids constitute 23 of the 35 residues of the C-terminal part of protein C, which is also palmitoylated on two cysteine residues. SP-B is formed from a large 381-residue precursor. The mature protein contains seven cysteines and disulfide bridges. Both proteins have major effects on the properties of the surfactant mixture. They promote rapid reorganization of lipid layers, an important consideration for the functioning of the surfactant. Infants lacking SP-B suffer severe respiratory failure with high mortality.[e]

The properties of the surfactant allow for rapid formation of a large area of lipid monolayer–air interface. The low surface tension and the ability to rapidly spread the mixture of lipids and proteins are essential.[o] At the end of the expiration stage of breathing the surfactant is present in the interface as a strong, tightly packed monolayer whose properties reflect the rigidity of the dipalmitoylphosphatidylcholine. The excess surfactant in the alveolar fluid forms liposome-like bilayer structures and also associates with proteins and calcium ions to form a lattice-like material called **tubular myelin**. Lipid in this form must be transferred rapidly into the air–liquid interface during inspiration.[p] The ability of phospholipids to pass through a hexagonal phase (Fig. 8-12) may also be important for this transition.[q,r]

One enzyme present in the surfactant fluid is an acid phosphatase able to hydrolyze phosphatidylglycerol phosphate, perhaps functioning in the final step of biosynthesis of the phosphatidylglycerol present in the surfactant.[c] Study of the action of the natural lung surfactant has led to development of artificial surfactant mixtures that are being used effectively to save many lives.[d]

---


[a] Rooney, S. A. (1979) *Trends Biochem. Sci.* **4**, 189–191

[b] Persson, A., Chang, D., Rust, K., Moxley, M., Longmore, W., and Crouch, E. (1989) *Biochemistry* **28**, 6361–6367

[c] Rooney, S. A., Young, S. L., and Mendelson, C. R. (1994) *FASEB J.* **8**, 957–967

[d] Jobe, A. H. (1993) *N. Engl. J. Med.* **328**, 861–868

[e] Kuroki, Y., and Voelkers, D. R. (1994) *J. Biol. Chem.* **269**, 25943–25946

[f] Shiffer, K., Hawgood, S., Haagsman, H. P., Benson, B., Clements, J. A., and Goerke, J. (1993) *Biochemistry* **32**, 590–597

[g] Pastrana-Rios, B., Flach, C. R., Brauner, J. W., Mautone, A. J., and Mendelsohn, R. (1994) *Biochemistry* **33**, 5121–5127

[h] Bangham, A. D. (1992) *Nature (London)* **359**, 110

[i] Rana, F. R., Harwood, J. S., Mautone, A. J., and Dluhy, R. A. (1993) *Biochemistry* **32**, 27–31

[j] McCormack, F. X., Calvert, H. M., Watson, P. A., Smith, D. L., Mason, R. J., and Voelker, D. R. (1994) *J. Biol. Chem.* **269**, 5833–5841

[k] Taneva, S., McEachren, T., Stewart, J., and Keough, K. M. W. (1995) *Biochemistry* **34**, 10279–10289

[l] Crouch, E., Persson, A., Chang, D., and Heuser, J. (1994) *J. Biol. Chem.* **269**, 17311–17319

[m] Ogasawara, Y., Kuroki, Y., and Akino, T. (1992) *J. Biol. Chem.* **267**, 21244–21249

[n] Korimilli, A., Gonzales, L. W., and Guttentag, S. H. (2000) *J. Biol. Chem.* **275**, 8672–8679

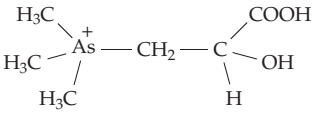
[o] Pastrana, B., Mautone, A. J., and Mendelsohn, R. (1991) *Biochemistry* **30**, 10058–10064

[p] Lipp, M. M., Lee, K. Y. C., Zasadzinski, J. A., and Waring, A. J. (1996) *Science* **273**, 1196–1199

[q] Perkins, W. R., Dause, R. B., Parente, R. A., Minchey, S. R., Neuman, K. C., Gruner, S. M., Taraschi, T. F., and Janoff, A. S. (1996) *Science* **273**, 330–332

[r] Discher, B. M., Maloney, K. M., Grainger, D. W., Sousa, C. A., and Hall, S. B. (1999) *Biochemistry* **38**, 374–383

resistant to phospholipase $A_1$ as well. These two properties appear to protect the naked cell membranes of the protozoa from their own phospholipases which may be secreted into the environment.[26] Marine algae form an arsenic-containing phospholipid *O*-phosphatidyltrimethylarsonium lactic acid.[27]

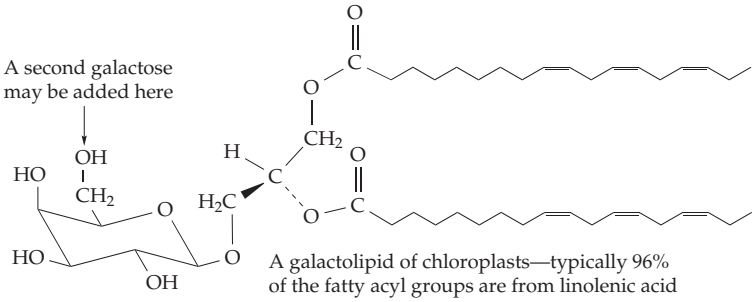


Trimethylarsonium lactic acid

They apparently do this as part of a scheme for detoxifying arsenate taken up with phosphate from the phosphate-poor ocean water.

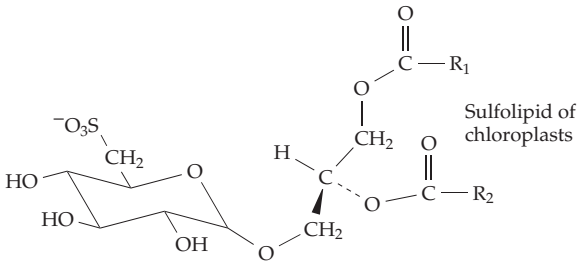
## 4. Glycolipids

The polar heads of the **glycoglycerolipids** lack phospho groups but contain *sugars* in glycosidic linkage.[28] Large amounts of the galactolipids shown in the following structure are found in chloroplasts.[29] The monogalactosyl diacylglycerol is said to be the most abundant polar lipid in nature.



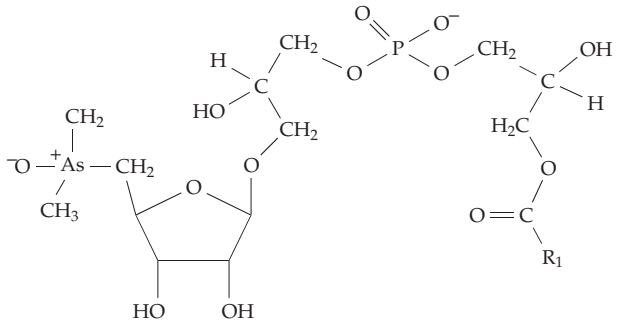

A galactolipid of chloroplasts—typically 96% of the fatty acyl groups are from linolenic acid

Chloroplasts also contain the following **sulfolipid**, an anionic sulfonate.



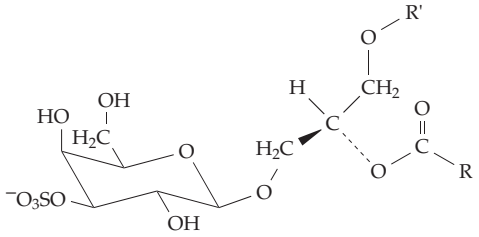

6-Sulfo-6-deoxy-α-D glucopyranosyl diglyceride

Marine algae as well as aquatic higher plants accumulate **arsenophospholipids**.[30]



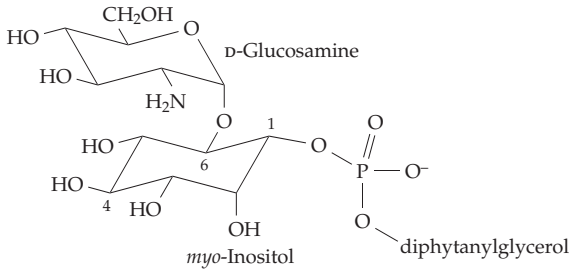
Arsenophospholipid of aquatic plants

The plasma membrane of mammalian male germ cells contains the following **sulfogalactosylglycerolipid**. It is found only in spermatozoa and testes, in which it accounts for 5–8% of total lipid, and in the brain, in which it accounts for only 0.2% of total lipid.[31]

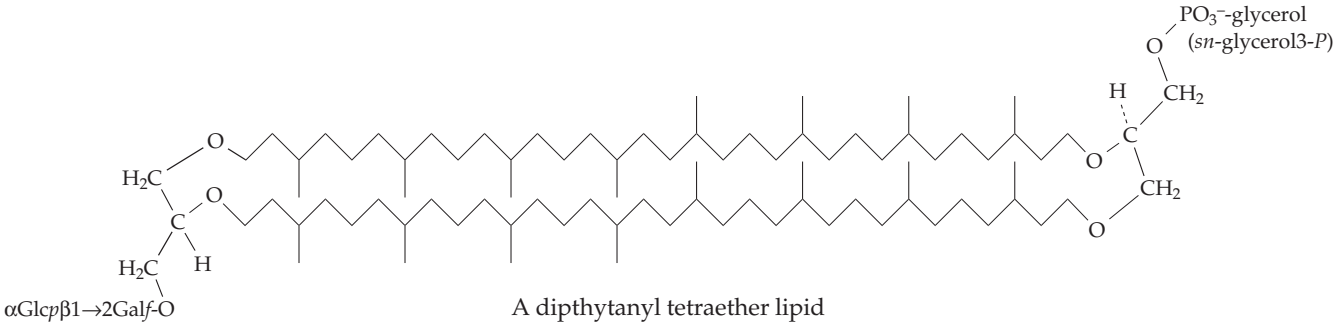


Sulfogalactosylglycerolipid, a sulfate ester

A variety of acylated glucolipids and phosphoglucolipids, including monoglucosyl and diglucosyl diacylglycerols, have been identified in membranes of the cell-wall-less bacterium *Acholeplasma laidlawii*.[32,32a] The following glycolipid from the methanogen *Methanosarcina*[33] is identical to the core structure of eukaryotic glycosylphosphatidylinositol membrane protein anchors (Fig. 8-13).




In addition to the previously mentioned phytanyl ether phospholipids, methanogens contain **diphytanyl tetraether lipids** that are both glycerophospholipids and glycolipids.

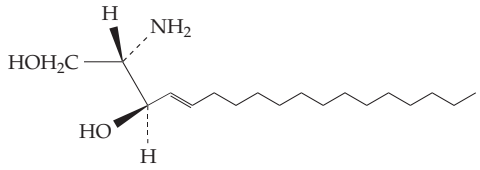



A dipthytanyl tetraether lipid

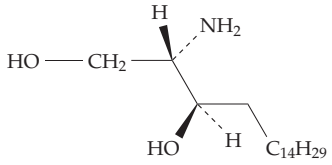
From a methanogenic bacterium.[21]

## 5. Sphingolipids

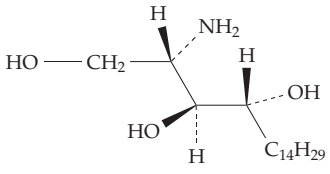
The backbone of the sphingolipids is the basic alcohol **sphingosine** (sphingenine) or a related long chain base. At least 60 such bases have been identified.[34,35] They vary in chain length from $C_{14}$ to $C_{22}$ and include members of the branched iso and anteiso series. Up to two double bonds may be present. Sphingosine contains 18 carbon atoms and is formed from palmitic acid and serine (Fig. 21-6). An intermediate in the formation of sphinganine is the saturated **sphinganine** (dihydrosphingosine), which is also a common component of animal sphingolipids. Hydroxylation of sphinganine to **phytosphingosine** occurs in both plants and animals, especially within glycolipids. The name comes from the fact that phytosphingosine was first discovered in plants.



Sphingosine (sphingenine)



Sphinganine (dihydrosphingosine)

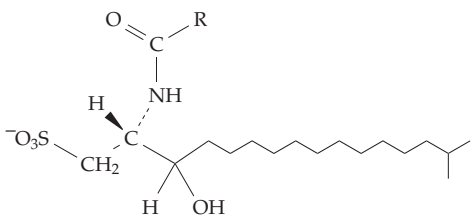


Phytosphingosine

Sphingosine-containing lipids are classified as **sphingophospholipids** (sphingomyelins) and **sphingoglycolipids**. In both cases the sphingosine is combined in amide linkage with a fatty acid to form a **ceramide** (Fig. 8-3) which still contains a free hydroxyl group able to combine with another component. In the sphingomyelins, which were first isolated from human brain by Thudicum in 1884, the additional component is usually phosphocholine (Fig. 8-3). Ceramide aminoethylphosphonates and related glycolipids occur in some invertebrates.[36,37]

The **cerebrosides** are glycosides of ceramide containing galactose or glucose. They are found in relatively large amounts in the brain where monogalactosylceramide predominates. Cerebrosides also occur in other animal tissues and to a lesser extent in plants. Many glycosphingolipids contain oligosaccharides of various sizes. When the oligosaccharide contains one or more residues of sialic acid the compound is known as a **ganglioside**.[38] Sialic acids are never found in plants. However, plants and fungi contain **phytoglycolipids**, which resemble gangliosides. Some contain inositol phosphates as well as sugars.[39–41] The structures of gangliosides may be very complex. Like glycoproteins, these substances are often located at the outer surface of cells where they may act as receptors for toxins, viruses, and hormones (see Section D).[42] Some of them carry attached blood group antigens (Box 4-C).[43,44] The sulfate esters of cerebrosides, known as **sulfatides** (Fig. 8-3), are also important components of membranes.[45]

Gliding bacteria such as *Capnocytophaga* contain sulfonolipids:[46]



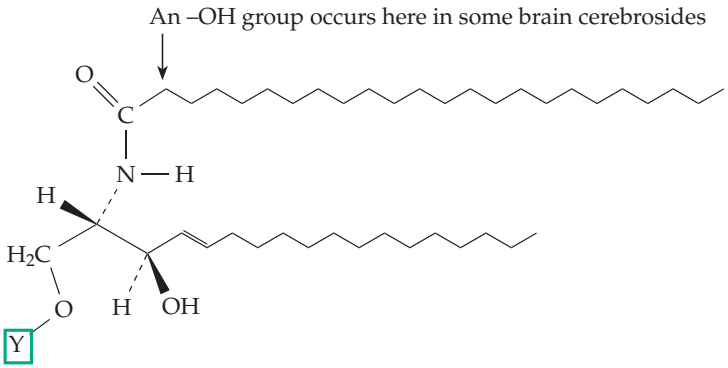
*N*-acylcapnine

A 1-deoxy-1-sulfonate analog of ceramides from gliding bacteria. R is a long-chain alkyl group.

These are analogs of the ceramides. The sulfonolipids seem to be necessary for the gliding movement of these bacteria across solid surfaces.[47]

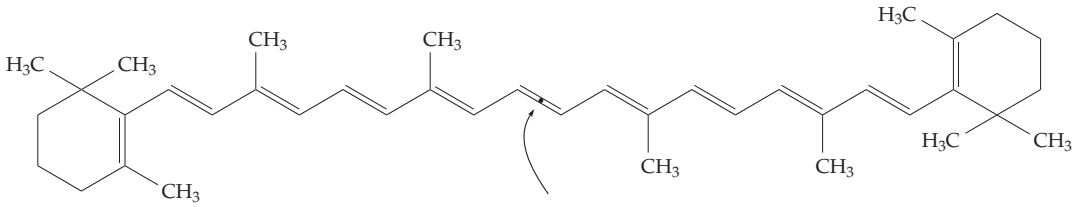
## 6. Sterols and Other Isoprenoid Lipids

A large group of isoprenoid lipids, including **sterols**, **terpenes**, and **carotenoid compounds**, are often present in membranes or in extracted lipids. Among these are the fat-soluble vitamins A, D, E, and K and the high polymers **rubber** and **gutta-percha**. The phytyl group of chlorophyll (Fig. 23-30), the phytanyl and diphytanyl groups of the lipids of methanogens, and the side chain of the pigment heme *a* (Fig. 16-5) are all related and are all derived from the precursor **prenyl pyrophosphate** (isopentenyl diphosphate). The whole group of compounds are often referred to as **polyprenyl** or by the older designation **isoprenoid**. The major discussion of polyprenyl compounds is found in Chapter 22 but the role of cholesterol and related compounds in membranes and the function of polyprenyl groups as membrane anchors for some proteins will be considered in this chapter.



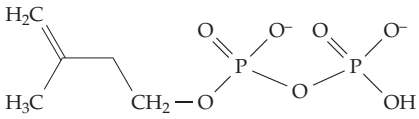

| | |
|---|---|
| Y = —H | The compounds in which Y = H are **ceramides**. Sphingolipids are often named as ceramide derivatives. |
| Y = —O—P(=O)(O⁻)—O—$CH_2CH_2$—$N^+(CH_3)_3$ (Phosphocholine) | **Sphingomyelins** |
| Y = —P(=O)(O⁻)—$CH_2CH_2NH_3^+$ | Ceramide aminoethyl phosphonates |
| Y = D-Galβ or D-Galβ1→4Glc– (The galactose bears a 3-sulfate group in cerebroside sulfatides, e.g., in lactosyl ceramide sulfate) | **Cerebrosides** or ceramide mono- and oligosaccharides |
| Y = GalNAc1→3Gal1→4Gal1→4Glc | Present in sphingolipid of red blood cell membranes |
| Y = —O—P(=O)(O⁻)—O—Inositol—Mannose | Present in **phytoglycolipid** of yeast |

**Figure 8-3** Structures of some sphingolipids.

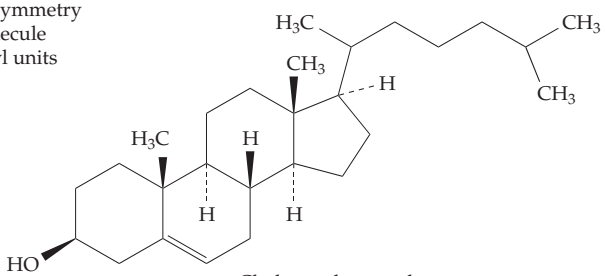


β-Carotene, a yellow pigment of carrots and other plants; converted in the human body to vitamin A.

Notice the center of symmetry in this 40-carbon molecule derived from 8 prenyl units



Prenyl diphosphate (pyrophosphate) (isopentenyl pyrophosphate), precursor of steroid,carotenoid, and other isoprenoid substances



Cholesterol, a sterol

## B. Membranes

In Chapter 7 we examined ways in which protein subunits can be stacked to form helices and closed oligomers. Another important arrangement of cell constituents is that of flat sheets or membranes.[48–54] Chemists, physicists, and biologists have mounted a sustained effort to understand these thin but remarkably tough outer surfaces of cells. However, consider the fact that a 7- to10-nm-thick plasma membrane of a cell of 10-μm diameter has less than 1/1000 the thickness of the cell and occupies only 0.5% of the total volume. The technical difficulties in studying such a membrane are great and are compounded by the fact that a cell contains more than one kind of membrane.

Membranes from many sources have been studied. One of these is **myelin**, the multilayered insulation that surrounds the axons of many nerve cells.[55–57] Myelin is derived from the plasma membrane of **Schwann cells** which lie adjacent to many neurons. Schwann cells literally wrap themselves around neuronal axons. Their cytoplasm is squeezed out leaving little but tightly packed membrane layers. Myelin membranes are the most stable known and also have the highest lipid content (80%). Another readily available experimental material is the plasma membrane of the human red bood cell, which can be prepared by osmotic rupture of the cells. The remaining **erythrocyte ghosts** contain ~1% of the dry matter of the cell and may have been studied more than any other membrane. A much investigated specialized membrane is the outer portion of the visual receptor cells known as **rods** (Chapter 23), which contains a closely packed and regular array of flat discs, each one consisting of a pair of membranes. Both membranes and cell walls of many kinds of bacteria have also been investigated.

### 1. The Structure of Membranes

Membranes consist largely of protein and lipid. The ratio (by weight) of protein to lipid varies from 0.25 in myelin to ~3.0 in bacterial membranes. In membranes of erythrocytes it is about 1.2 and a ratio of about 1.0 may be regarded as typical for animal cells. Small amounts of carbohydrates (<5%) are present, as are traces of RNA (<0.1%).

In 1926, Gorter and Grendel calculated that the erythrocyte ghost contained just enough lipid to form a 3.0- to 4.0-nm-thick layer around the cell. Apparently they reached this correct conclusion only because their measurements of pressures of surface films contained compensating errors.[58] Nevertheless, this information, together with the known propensity of lipids to aggregate in **micelles** in which the hydrocarbon "tails" clustered together and the polar "heads" protruded into the surrounding water,[59] led Danielli and Davsen in 1935 to propose the **lipid bilayer** structure for membranes.[60] Its essential features are indicated in Fig. 8-4. Hydrophobic bonding holds the extended hydrocarbon chains together, while the polar groups of the phospholipid molecules may interact with proteins on the sides of the bilayer. The original proposal assumed an extended β structure for the proteins, which would allow them to coat the bilayer uniformly on both sides. However, this is not correct. Proteins are sometimes embedded in the bilayer, sometimes protrude through it, and sometimes are attached on one surface, most often the cytoplasmic surface of the plasma membrane. These concepts were brought together by Singer and Nicolson in 1972 in the **fluid mosaic model** of membrane structure[61] (Fig. 8-5). The lipid bilayer still provides the basic structure upon which the complex membranes of living organisms are assembled.[52,54,62–65] The term "fluid" refers to the fact

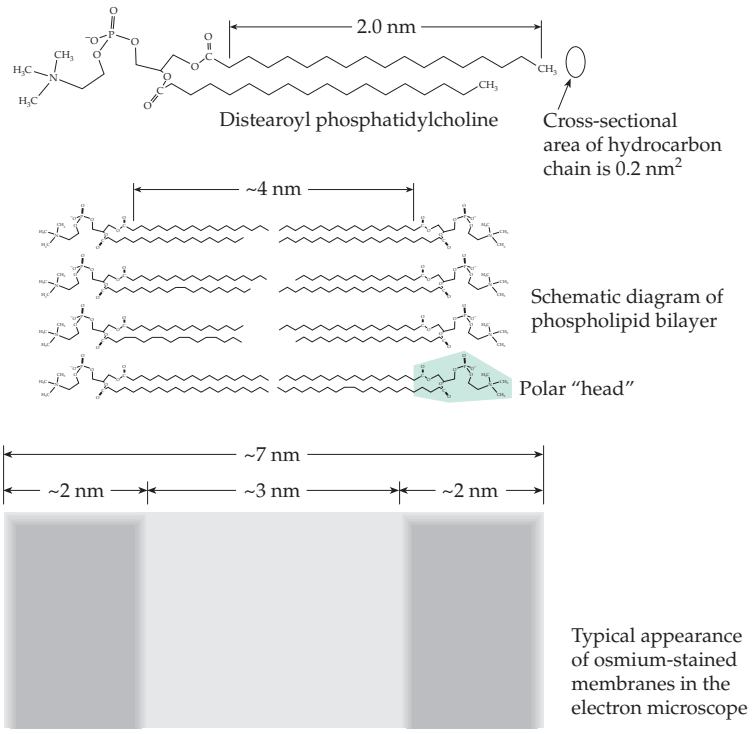



**Figure 8-4** Bimolecular lipid layers and membranes. (Top) A molecule of phosphatidylcholine. (Center) Lipid bilayer structure. (Bottom) Bilayer structure as seen by the electron microscope with osmium tetroxide staining.

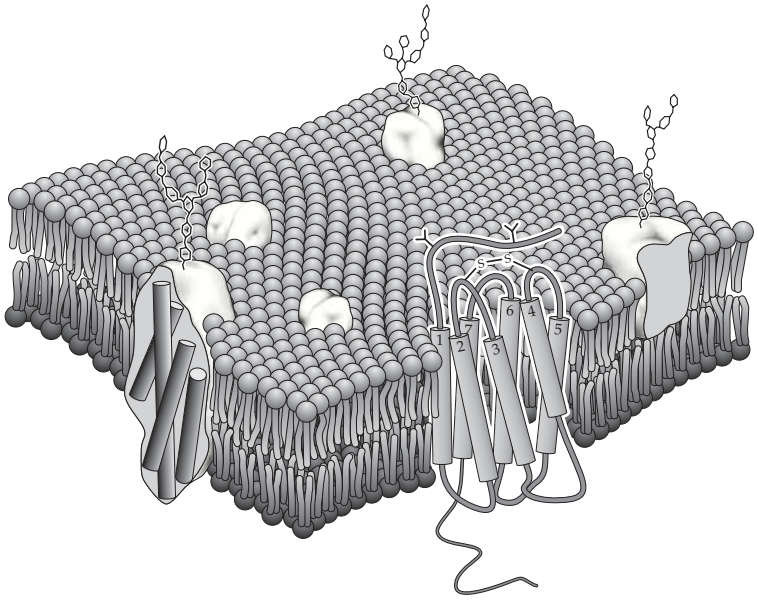



**Figure 8-5** The fluid mosaic model of Singer and Nicolson.[61] Some integral membrane proteins, which are shown as irregular solids, are dissolved in the bilayer. Transmembrane proteins protrude from both sides. One of these is pictured as a seven-helix protein, a common type of receptor for hormones and for light absorption by visual pigments. Other proteins adhere to either the outer or the inner surface. Many membrane proteins carry complex oligosaccharide groups which protrude from the outer surface (Chapter 4). A few of these are indicated here as chains of sugar rings.

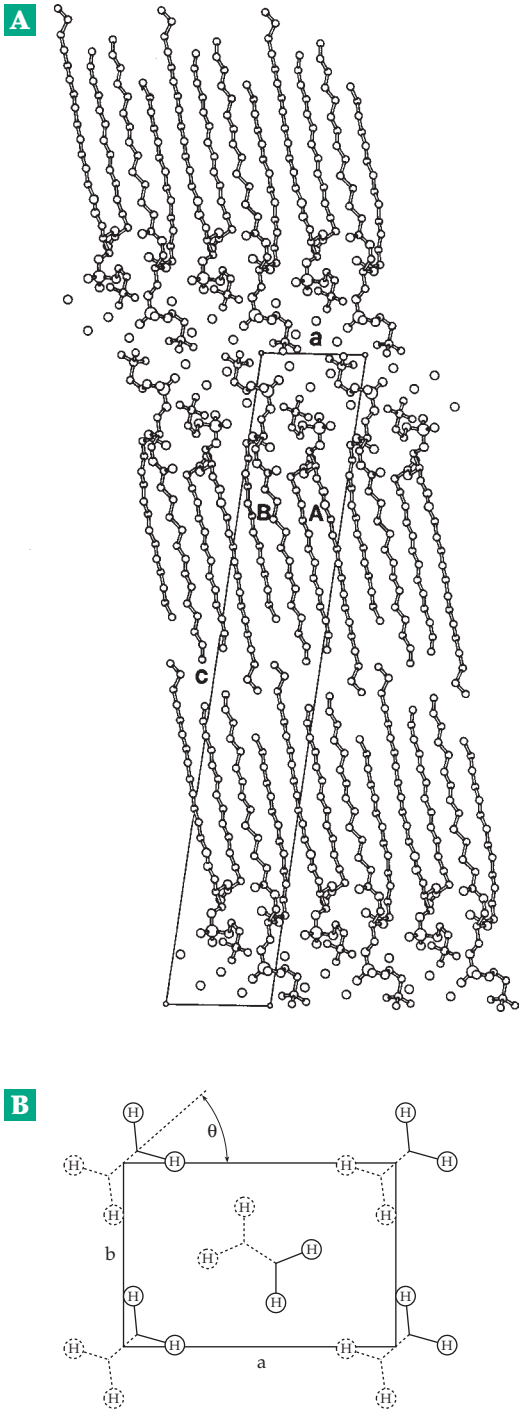



**Figure 8-6** (A) Molecular packing of 2,3-dimyristoyl-D-glycero-1-phosphocholine dihydrate. The two molecules in the asymmetric unit are labeled 1 and 2. The position of the water molecules is indicated either by W1–W4 or by small open circles. Hydrogen bonds are represented by dotted lines. From Pascher *et al.*[66] (B) Two-dimensional "orthorhombic" arrangement of hydrocarbon chains in a crystalline alkane. The $a-b$ plane corresponds to the plane of the bilayer surface; the long axes of the acyl chains project from the page. From Cameron *et al.*[67]

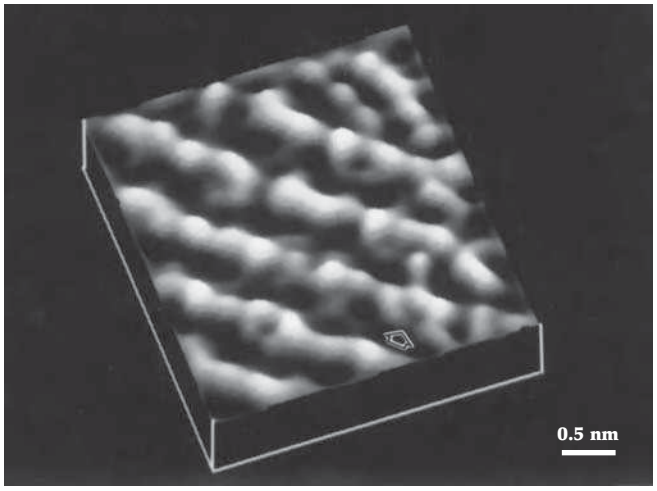



**Figure 8-7** Atomic force microscope image of a dimyristoylphosphatidylethanolamine bilayer deposited by the Langmuir–Blodgett technique (see Fig. 8-8) at a specific molecular area of 0.41 $nm^2$ and a surface pressure of 40 mN/m on a freshly cleaved mica substrate. The images were taken under water. The long, uniformly spaced rows are roughly 0.7–0.9 nm in spacing. The modulation along the rows, with rounded bright spots roughly every 0.5 nm, corresponds to the individual headgroups of the phosphatidylethanolamine molecules. The lattice spacing is identical to that measured by X-ray diffraction at the air–water interface. The area per molecule in the AFM image is ~0.4 $nm^2$. From Zasadzinski *et al.*[68]

that at temperatures suitable for growth and metabolism the hydrocarbon chains are not rigidly packed in the center of the bilayer but are "molten" (see Section 2). However, at a low enough temperature they become rigid and pack together in a manner similar to that of the chains in crystals of the phosphatidylcholine shown in Fig. 8-6A. These crystals consist of stacked bilayers of thickness 5.5 nm.[66,69] The scanning tunneling and atomic force microscopes have provided direct views of a similar arrangement of side chains in a monomolecular fatty acid layer[68,70,71] (Fig. 8-7). Measurements on multilamellar vesicles of dipalmitoylphosphatidylcholine give bilayer thicknessess from 5.4 nm for dehydrated vesicles to 6.7 nm for the biologically relevant fully hydrated bilayers.[72]

***Lipids of membranes.*** Approximately 1500 different lipids have been identified in the myelin of the central nervous system of humans. About 30 of these are present in substantial amounts.[73] The distribution of the different lipids varies markedly between membranes from different sources (Table 8-3) making generalization difficult. However, phospholipids are apparently always present and, except in chloroplasts, make up from 40% to over 90% of the total lipid (Table 8-3).

Five kinds of phospholipid predominate: phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, phosphatidylglycerols, and sphingomyelin. Usually there are also small amounts of phosphatidylinositol. The major phospholipid in animal cells is phosphatidylcholine, but in bacteria it is phosphatidylethanolamine. The phospholipids of *E. coli* consist of 80% phosphatidylethanolamine, 15% phosphatidylglycerol, and 5% diphosphatidylglycerol (cardiolipin). Significant amounts of cardiolipin are found only in bacteria and in the inner membrane of mitochondria. Sphingomyelin is almost absent from mitochondria, endoplasmic reticulum, or nuclear membranes.

Glycolipids are important constituents of the plasma membranes, of the endoplasmic reticulum, and of chloroplasts. The cerebrosides and their sulfate esters, the sulfatides, are especially abundant in myelin. In plant membranes, the predominant lipids are the galactosyl diglycerides.[29,74] The previously described ether phospholipids (archaebacteria), ceramide aminoethylphosphonate (invertebrates), and sulfolipid (chloroplasts) are also important membrane components.

Cholesterol makes up 17% of myelin and is present in plasma membranes. However, it usually does not occur in bacteria and is present only in trace amounts in mitochondria. Related sterols are present in plant membranes. Esters of sterols occur as transport forms but are not found in membranes. Membrane bilayers, likewise, contain little or no triacylglycerols, the latter being found largely as droplets in the cytoplasm.

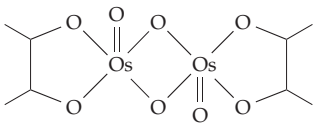
Quantitatively minor membrane components with important biological functions include **ubiquinone**, which is present in the inner mitochondrial membrane, and the **tocopherols**. Plant chloroplast membranes contain chlorophyll, carotenes, and other lipid-soluble pigments.

***Liquid crystals, liposomes, and artificial membranes.*** Phospholipids dissolve in water to form true solutions only at very low concentrations (~$10^{-10}$ M for distearoyl phosphatidylcholine). At higher concentrations they exist in **liquid crystalline phases** in which the molecules are partially oriented. Phosphatidylcholines (lecithins) exist almost exclusively in a **lamellar** (smectic) phase in which the molecules form bilayers. In a warm phosphatidylcholine–water mixture containing at least 30% water by weight the phospholipid forms **multilamellar vesicles**, one lipid bilayer surrounding another in an "onion skin" structure. When such vesicles are subjected to ultrasonic vibration they break up, forming some very small vesicles of diameter down to 25 nm which are surrounded by a single bilayer. These unilamellar vesicles are often used for study of the properties of bilayers. Vesicles of both types are often called **liposomes**.[75–77]

When liposomes are stained with osmium tetroxide or potassium permanganate, embedded, and sectioned

for electron microscopy, their membranes show a characteristic three layered structure similar to that observed for biological membranes. Two darkly stained lines ~2–2.5 nm thick are separated by a clear space ~2.5–3.5 nm wide in the center. Both myelin and the retinal rod outer segments show closely spaced pairs of such membranes with a combined width of 18 nm. These results seemed to support the original Davsen–Danielli model. However, many questions must be raised about the interpretation of these results. Why does $OsO_4$ stain only the outer protein layer when it is known to react also with double bonds of hydrocarbon side chains of lipids to form osmate esters which are readily reduced to a diol and osmium?[77a-c,78] Membranes from which most of the lipid has been extracted still stain with $OsO_4$ to give three-layered electron micrographs. Perhaps little can be concluded from the three-layered appearance. We have learned that it is difficult to determine even the thickness, let alone the complete structure of an object that is only 6–10 nm thick.

Strong support for the lipid bilayer model comes from the preparation of another type of artificial membrane which can be made from a solution of phosphatidylcholine or of a mixture of phospholipids plus cholesterol in a hydrocarbon solvent. A droplet of solution is placed on a small orifice in a plastic sheet, separating two compartments filled with an aqueous medium (Fig. 8-8). The solution in the orifice quickly drains, just as does a soap bubble, and the resulting film eventually becomes so thin that the bright colors disappear and a **"black membrane"** is formed. Similar membranes, but without a residual content of hydrocarbon solvent, have been created by apposition of two lipid monolayers formed at an air–water interface.[79,80] The thickness of such artificial membranes is thought to be only 6–9 nm. Resilient and self-sealing, the membranes can be stained with $OsO_4$ to give a typical three-layered pattern.



An osmate ester formed from two unsaturated groups

**TABLE 8-3**
**Estimated Chemical Compositions of Some Membranes**

| Compound | Percentage of total dry weight of membrane[a] | | | | | |
|---|---|---|---|---|---|---|
| | Myelin (bovine) | Retinal rod | Plasma membrane (human erythrocyte) | Mitochondrial membranes | *E. coli*[b,c,d] (inner and outer membranes) | Chloroplasts[e] |
| Protein | 22 | 59 | 60 | 76 | 75 | 48 |
| Total lipid | 78 | 41 | 40 | 24 | 25 | 52[f] |
| Phosphatidylcholine | 7.5 | 13 | 6.9 | 8.8 | | |
| Phosphatidylethanolamine | 11.7 | 6.5 | 6.5 | 8.4 | 18 | |
| Phosphatidylserine | 7.1 | 2.5 | 3.1 | | | |
| Phosphatidylinositol | 0.6 | 0.4 | 0.3 | 0.75 | | |
| Phosphatidylglycerol | | | | | 4 | |
| Cardiolipin[g] | | 0.4 | | 4.3 | 3 | |
| Sphingomyelin | 6.4 | 0.5 | 6.5 | | | |
| Glycolipid | 22.0 | 9.5 | Trace | Trace | | 23 |
| Cholesterol | 17.0 | 2.0 | 9.2 | 0.24 | | |
| Total phospholipid | 33 | 27 | 24 | 22.5 | 25 | 4.7 |
| Phospholipid as a percentage of total lipid | 42 | 66 | 60 | 94 | >90% | 9 |

[a] Dewey, M. M. and Barr, L. (1970) *Curr. Top. Membr. Transp.* **1**, 6.
[b] Kaback, H. R. (1970) *Curr. Top. Membr. Transp.* **1**, 35–99.
[c] Mizushima, S. and Yamada, H. (1975) *Biochim. Biophys. Acta.* **375**, 44–53 .
[d] Yamato, I. Anraku, Y. and Hirosawa, H. (1975) *J. Biochem. (Tokyo)* **77**, 705–718. These investigators found 67% protein, 21% lipids, 10% carbohydrate, and 2% RNA.
[e] Lichtenthaler, H. K. and Park, R. B. (1963) *Nature (London)* **198**, 1070–1072.
[f] About 14% is accounted for by chlorophyll, carotenoids, and quinones.[e]
[g] Diphosphatidylglycerol (Fig. 8-2).

The study of monolayers formed on a water surface has also provided important information. A thin film of an amphiphilic (containing both polar and nonpolar groups) compound such as a fatty acid is prepared. This is done by depositing a small quantity of the compound dissolved in a volatile solvent on a clean aqueous surface between the barriers of a **Langmuir trough** (Fig. 8-8).[81,82] The difference in surface tension ($\pi$) across the barriers is measured with a suitable device[81] for different areas of the monolayer, i.e., for different positions of the moveable barrier. The value of $\pi$ is low for expanded monolayers and falls to nearly zero when the surface is no longer completely covered. The pressure reaches a plateau when a compact monolayer is formed, after which it rises again (Fig. 8-8B). At very high values of $\pi$ the monolayer collapses (buckles). Both the cross-sectional area per molecule in the monolayer and the collapse pressure can be determined. For typical fatty acids, regardless of chain length, the area covered is only ~0.2 nm$^2$ per molecule indicating that the fatty acid chains are stacked vertically to the surface in the monolayer. The collapse pressure is higher for longer molecules as a result of the greater number of van der Waals interaction between the chains. **Langmuir–Blodgett layers** are prepared by transferring one or more monolayers onto a smooth solid surface (Fig. 8-7).[82,83]

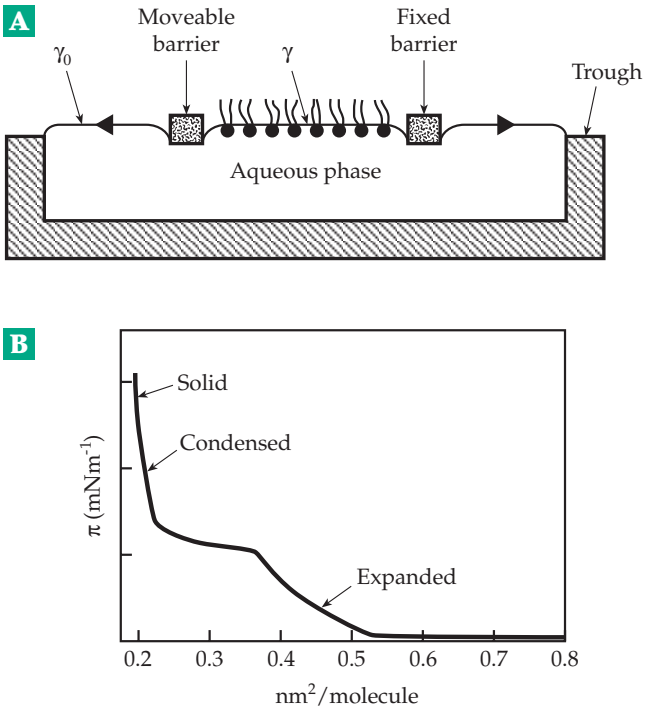
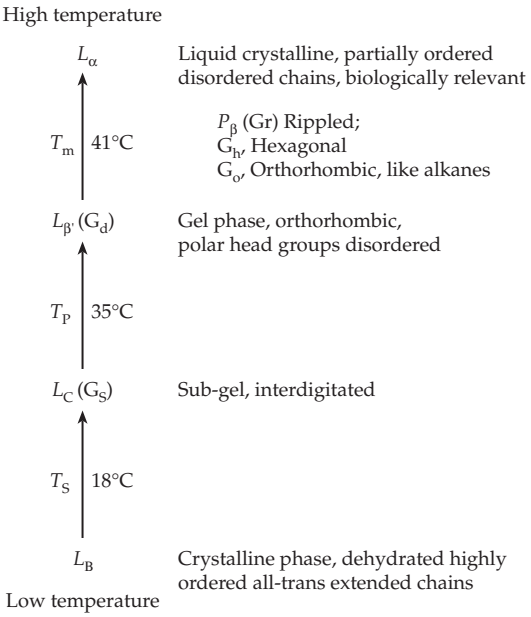



**Figure 8-8** (A) The Langmuir–Adam film balance. Tension on the moveable barrier is recorded for different areas of the surface between the barriers. This gives the surface pressure $\pi$, which is the difference between the surface tension ($\gamma_o$) of a clean aqueous surface and that of a spread monolayer ($\gamma$): $\pi = \gamma_o - \gamma$. Courtesy of Jones and Chapman.[81] (B) Surface pressure ($\pi$)–area per molecule isotherm for a typical fatty acid (e.g., pentadecanoic acid $C_{14}H_{29}CO_2H$) at the aqueous–air interface. From Knobler.[81a]

***Physical properties of membrane lipids.*** A completely extended $C_{18}$ fatty acid chain as shown in Fig. 8-4 has a length of ~2.0 nm and occupies, either in crystals or in monolayers, when viewed "end-on," an area of ~0.2 nm$^2$. The hydrocarbon layer in a lipid bilayer containing such chains would have a thickness of about 4.0 nm; that determined by X-ray diffraction for myelin is ~3.5 nm. However, for artificial black membranes the thickness of the hydrocarbon layer can be as little as 3.1 nm when all solvent is removed.[84] These and many other results[85] indicate that the hydrocarbon chains are to some extent folded and that the membrane is expanded over that expected according to the simplest model.

Structure determinations on crystalline alkanes confirm that the chains exist in a completely extended conformation and that adjacent chains often pack together in the orthorhombic arrangement shown in Fig. 8-6B. As the temperature of such crystals is raised a series of solid–solid phase transitions is observed below the melting point of the crystals.[86] These can be detected by changes in the infrared absorption spectrum or by small amounts of heat absorption revealed by **differential scanning calorimetry** (Fig. 8-9). Each new phase permits a greater degree of mobility for the hydrocarbon chains. Thus, at a high enough temperature but below the melting point, the chains are able to rotate freely about their own axes in a so-called **hexagonal phase**. Now the chains are packed in a hexagonal array instead of the orthorhombic array of Fig. 8-6B. At intermediate temperatures, some of the chains may assume nonplanar conformations and changes in the tilt of the hydrocarbon chains (Fig. 8-6) may occur.

Similar phase transitions are observed for bilayers.[88–90] For dipalmitoyl phosphatidylcholine the first detectable **subtransition**[91] is centered at a temperature $T_s$ of 18°C. The second, known as the **pretransition**, occurs at 35°C ($T_p$). The structure below $T_s$ may be described as rigid or crystalline and that above $T_s$ as a **gel** in which the hydrocarbon side chains twist and turn much more freely but in which the orthorhombic packing is maintained.[86] Above $T_p$ the head groups become disordered. Although the orthorhombic packing of the tails may be maintained, there are several distinct phases,[92,93] including one or more in which the gel is thought to assume a structure analogous to that in the hexagonal phase of hydrocarbons. At 41°C the **main transition** occurs.




This is a sharper transition with a well-defined melting temperature designated $T_m$. Above $T_m$ the lipid is in the lamellar liquid crystalline or $L_\alpha$ state. The bilayer continues to hold together, but the fatty acid chains have melted and are now free to rotate and undergo twisting movements more freely than at lower temperatures (Fig. 8-11). The main transition is highly, but not completely, cooperative. Thus, the melting of the membrane occurs over a range of several degrees. The presence in biological membranes of a variety of different components containing a variety of fatty acid chains leads to a broadening of the melting range.

The behavior of bilayers is strongly influenced by the lipid composition. Phospholipids containing saturated, long-chain fatty acids have high transition temperatures. The presence of unsaturated fatty acyl groups with *cis* double bonds in membrane lipids encourages folding of the hydrocarbon chains and lowers $T_m$. Even a single double bond lowers $T_m$, the decrease being greatest when the double bond is near the center of the chain.[94–96] While $T_m$ for dipalmitoyl phosphatidylcholine is 41°C, that of 1,2-dipalmitoyl phosphatidyl-*sn*-glycerol, which lacks the phosphocholine head group, is 70°C. This falls to 11.6°C for the polyunsaturated 1-stereoyl-2-linoleoyl-*sn*-glycerol, whose melting curve is shown in Fig. 8-9.[87] This lipid also shows a complex phase behavior and a melting point for the stable, crystalline β′ phase higher than that of the α phase.

Inclusion of other molecules of irregular shape within membranes also lowers $T_m$. However, a molecule of cholesterol can pack into a bilayer with a cross-sectional area of 0.39 nm$^2$, just equal to that of two hydrocarbon chains.[49] It tends to harden membranes above $T_m$ but increases mobility of hydrocarbon chains below $T_m$.[97–100] A complex of cholesterol and phosphatidylcholine may form a separate phase within the membrane.[101,102] The ether-linked plasmalogens may account for over 30% of the phosphoglycerides of the white matter of the brain and of heart and ether linked phospholipids are the major lipids of many anaerobic bacteria.[103] Their $T_m$ values are a few degrees higher than those of the corresponding acyl phospholipids.[104]

Between the pretransition temperature and $T_m$ solid and liquid regions may coexist within a bilayer.[101] The term **lateral phase separation** has been applied to this phenomenon.[105,106] Since changes in the equilibrium between solid and liquid can be induced readily, e.g., by changes in the ionic environment surrounding the bilayer, lateral phase separation may be of significance in such phenomena as nerve conduction.[107]

The phase transitions in bilayers can be recognized in many ways. Differential scanning colorimetry has already been mentioned. Another approach is to measure the spacing between molecules by X-ray diffraction. The cross-sectional area occupied by a phospholipid in a bilayer is always greater than the 0.40 nm$^2$ expected for closest packing of a pair of extended hydrocarbon chains.[39,85]

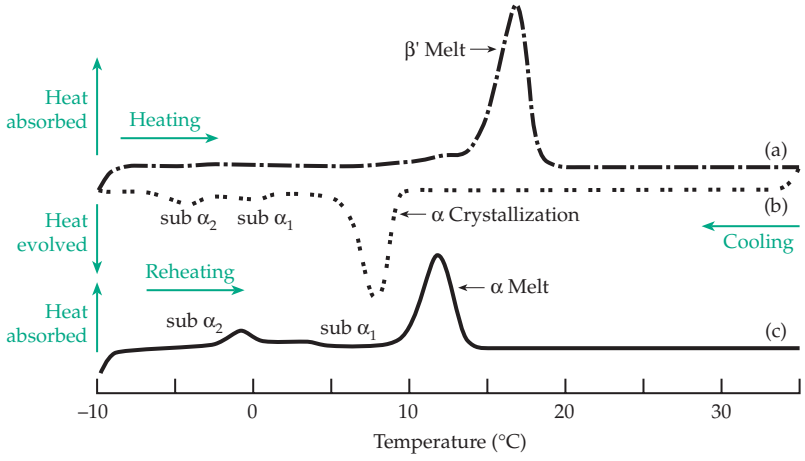



**Figure 8-9** Differential scanning calorimetric curves for 1-stearoyl-2-linoleoyl-*sn*-glycerol. (A) Crystals of the compound grown from a hexane solution were heated from –10° to 35°C at a rate of 5°C per minute and the heat absorbed by the sample was recorded. (B) The molten lipid was cooled from 35° to –10°C at a rate of 5° per minute and the heat evolved was recorded as the lipid crystallized in the α phase and was then transformed through two sub-α phases. (C) The solid was reheated. From Di and Small.[87] Courtesy of Donald M. Small.

Below $T_m$ the spacing between chains is about 0.42 nm corresponding to close packing of the fatty acid chains in a hexagonal array with an area per phospholipid of 0.41 nm$^2$. As the temperature is raised above $T_m$ the spacing increases[85] to give an average area per phospholipid of 0.64–0.73 nm$^2$. Another technique (Box 8-C) is to study a **spin label** by EPR while yet another is to observe the fluorescence of a **polarity-dependent fluorescence probe** such as *N*-phenylnaphthylamine or other fluorescent probes[108] (see Chapter 23). The compound is incorporated into the membrane and undergoes changes in the intensity of its fluorescence when the state of the membrane is altered.

A variety of NMR techniques are being applied[109–113] both to liposomes and to natural membranes.[111,114] Incorporation of $^{13}C$ or $^2H$ into various positions in the hydrocarbon chains has allowed measurements of the relative degree of mobility of the chains at different depths in the bilayer (Fig. 8-10).[109,115–117] The results are in agreement with statistical mechanical predictions that configurational freedom increases with depth toward the midplane of the bilayer. Separation of a bilayer into two or more phases can be observed using $^2H$- or $^{31}P$- NMR.[118–120] The orientation and dynamic behavior of various head groups has been explored,[110,121] as have effects of mixing into the bilayer other lipids such as glycosphingolipids[122] and cholesterol.[123,124] Crystalline phospholipids are being investigated by solid-state NMR.[125]

Fourier transform infrared spectroscopy[126,127] also provides information about conformation of both hydrocarbon chains and head groups. EPR spectroscopy (Box 8-C) with doxyl probes on carbon atoms at different depths within the bilayer has also been employed.[128]

In recent years **molecular dynamics simulations** have been used to predict behavior of membranes. As is indicated in Fig. 8-11, the molten interior of the liquid crystalline $L_\alpha$ state is portrayed clearly.[129–131] In the gel state the hydrocarbon chains maintain a closer packing and undergo coordinated movement.[88] It is difficult to know how realistic the simulations are. To calibrate the method efforts are made to correctly predict a series of known properties such as density and area per lipid (0.61 nm).[130]

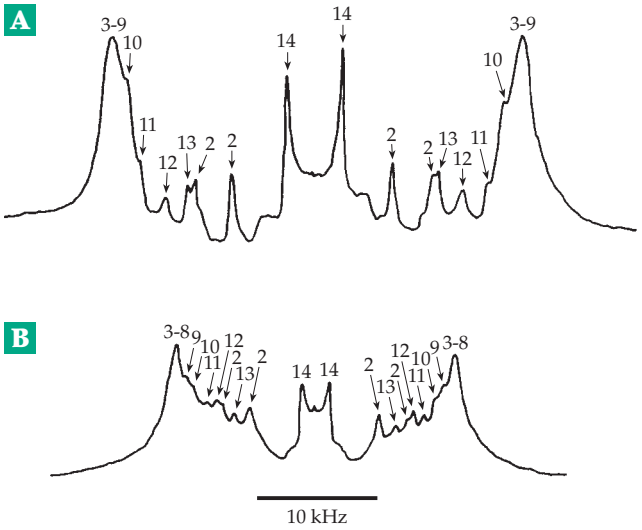
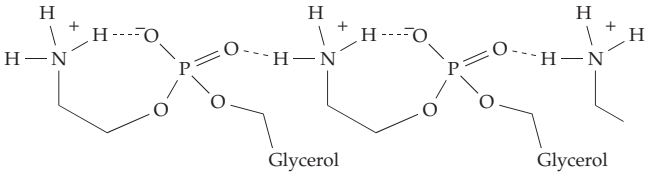



**Figure 8-10** $^2H$ NMR spectra of dimyristoyl phosphatidylcholine-$d_{27}$/water in lamellar phases at 40°C. One chain of the phosphatidylcholine is fully deuterated, containing 27 atoms of $^2H$. The mole ratios of water to lipid were 5.0 in (A) and 25.0 in (B). The average interfacial areas per alkyl chain as measured by X-ray diffraction were 0.252 nm$^2$ for (A) and 0.313 nm$^2$ for (B). $^2H$ NMR spectra are presented as "powder patterns" because the lipid molecules are randomly oriented in the magnetic field of the spectrometer as if in a powder. This gives rise to pairs of peaks symmetrically located on both sides of the origin. The separation distances are a measure of the quadrupole splitting of the NMR absorption line caused by the $^2H$ nucleus. The various splittings of the resonances of the 13 $-CH_2-$ and one $-CH_3$ groups reflect differences in mobility.[109] The peaks have been assigned tentatively as indicated. From Boden, Jones, and Sixl.[115] Courtesy of N. Boden.

***Functions of phospholipid head groups.*** The dipolar ionic head groups of phosphatidylcholine and phosphatidylethanolamine occupy about the same cross-sectional area as the two hydrocarbon tails. Thus, they are in rather close contact with each other. In crystals chains of hydrogen-bonded atoms may be formed. In phosphatidylethanolamine the phosphate and $-NH_3^+$ ions may alternate in these chains.[132]




In phosphatidylcholine, in which the nitrogen is surrounded by methyl groups and cannot form this kind of chain, water molecules bridge between the phosphates but the positive charges still interact with the adjacent negative charges.

The chains of hydrogen bonds between the head groups of phosphatidylethanolamine help to stabilize the bilayer and are apparently responsible for the elevation of $T_m$ by 10–30° above that observed for phosphatidylcholine.[132] In contrast, the negatively charged carboxyl groups of phosphatidylserine make the membrane less stable. The melting point is increased if the pH is lowered, protonating these groups. Their presence also makes the membrane sensitive to the concentration of cations.[133] The same is true of phosphatidylglycerol, whose head group contains a negatively charged phosphate without an attached

counterion. Addition of calcium ions increases $T_m$ greatly and causes either phosphatidyglycerol or phosphatidylserine to form a separate phase with a more crystalline-like packing of the hydrocarbon side chains.[134] Hydrogen bonding between head groups also occurs with glycolipids.[135]

***Non-bilayer structures of phospholipids.*** Under appropriate conditions some aqueous phospholipids can exist in non-bilayer phases, a fact that may be of considerable biological importance.[119,136,137] In the presence of $Ca^{2+}$ some pure phospholipids can be converted to the **inverted hexagonal** or $H_{II}$ phase (Fig. 8-12).[136,138–140] In this phase the phospholipid heads are clustered together in cylindrical "inverted" micelles which pack in a hexagonal array. The ease with which this transition can occur is increased by the presence of small amounts of diacylglycerols or lysolecithins.[141] Some lipids, such as the galactosyldiacylglycerol of chloroplasts, do not form bilayers but prefer the hexagonal phase structure.[29,32a] This is thought to be a result of very high curvature of a bilayer that arises from the sizes and packing of their head groups. Another phase, even though it is liquid, has a three-dimensional cubic symmetry.[142–144a] It apparently consists of a complex arrangement of polyhedral bilayer surfaces with interpenetrating water channels between them.[143]

***Membrane fluidity and life.*** In agreement with the known behavior of bilayers, the lipids of most membranes in all organisms are partially liquid at those temperatures suitable for life. Organisms have developed at least three distinct means of ensuring that membrane lipids remain liquid.[145] (1) In our bodies (as well as in *E. coli*) the unsaturated fatty acids that are present lower the melting point. Mutants of *E. coli* that are unable to synthesize unsaturated fatty acids cannot live unless these materials are supplied in the medium.[146] (2) In *Bacillus subtilis*, which contains no unsaturated fatty acids when grown at 37°C, and in other gram-positive bacteria, more than 70% of membrane

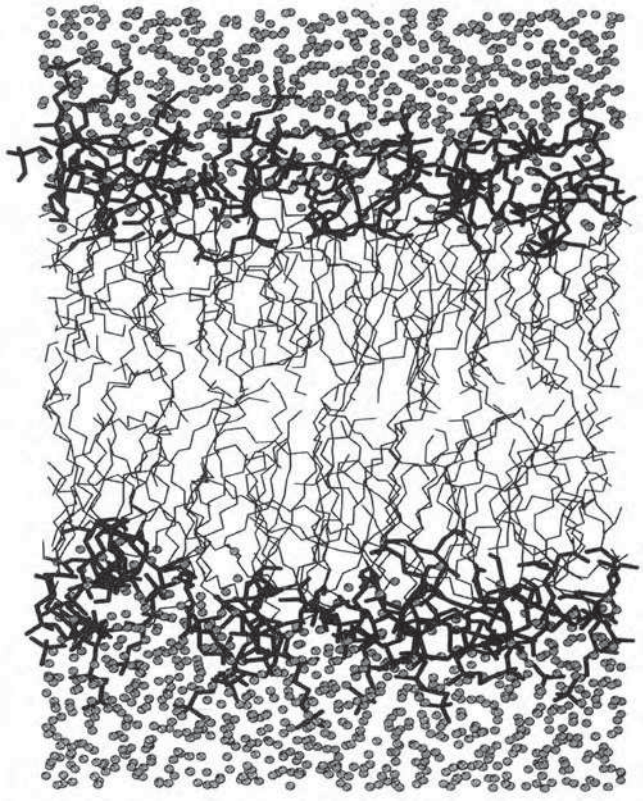


**Figure 8-11** Results of simulated motion in a lipid bilayer consisting of 64 molecules of dipalmitoylphosphatidylcholine and 23 water molecules per lipid at a pressure of 2 atm and 50°C. The view is that observed after 500 ps of simulation. Bold lines represent the head group and glycerol parts of the structures and the thin lines the hydrocarbon chains. The gray spheres represent water molecules. From Berger, Edholm, and Jähnig.[130] Courtesy of Dr. Olle Edholm.

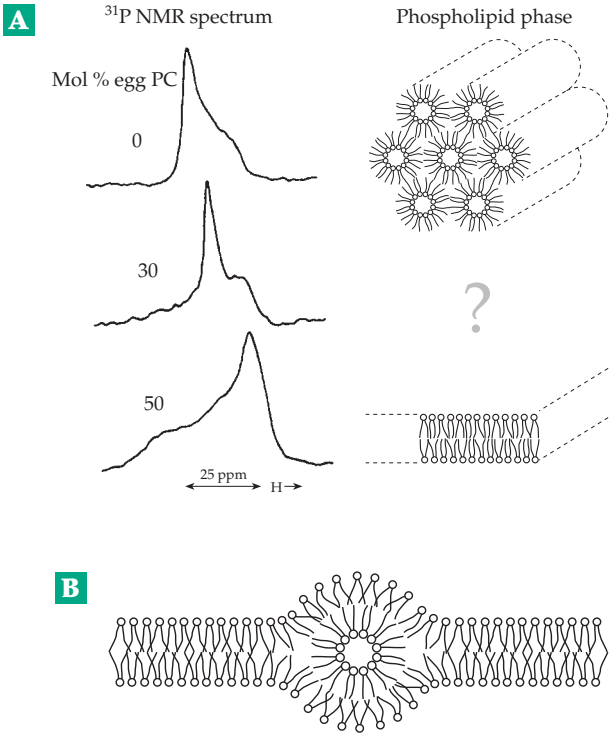



**Figure 8-12** (A) $^{31}P$ NMR spectra of different phospholipid phases. Hydrated soya phosphatidylethanolamine adopts the hexagonal $H_{II}$ phase at 30°C. In the presence of 50 mol% of egg phosphatidylcholine only the bilayer phase is observed. At intermediate (30%) phosphatidylcholine concentrations an isotropic component appears in the spectrum. (B) Inverted micelles proposed to explain "lipidic particles" seen in freeze fracture micrographs of bilayer mixture of phospholipids, e.g., of phosphatidylethanolanine + phosphatidylcholine + cholesterol. From de Kruijft *et al.*[119] Courtesy of B. de Kruijft.

## BOX 8-C ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTRA AND "SPIN LABELS"

Unpaired electrons have magnetic moments and are therefore suitable objects for magnetic resonance spectroscopy. The technique is similar to NMR spectroscopy, but microwave frequencies of ~$10^{10}$ Hz are employed, the energies being ~100 times greater than those used in NMR.[a,b] Unpaired electrons are found in organic free radicals and in certain transition metal ions, both of which are important to many enzymatic processes. Furthermore, **spin labels** in the form of stable organic free radicals, can be attached to macromolecules at many different points.[a,c–e] Coupling of such artificially introduced unpaired electrons with the magnetic moments of other unpaired electrons or of magnetic nuclei can often be observed by EPR techniques.

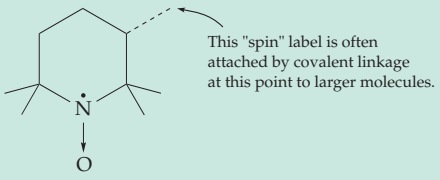
The conditions for absorption of energy in the EPR spectrometer are given by the equation

$$h\nu = g\beta H_o$$

which is identical in form to that for NMR spectroscopy. Here $H_o$ is the external magnetic field strength and β is a constant called the Bohr magneton. The value of $g$, the **spectroscopic splitting factor**, is one of the major characteristics needed to describe an EPR spectrum. The value of $g$ is exactly 2.000 for a free electron but may be somewhat different in radicals and substantially different in transition metals. One factor that causes $g$ values to vary with environment is **spin-orbit coupling** which arises because the $p$ and $d$ orbitals of atoms have directional character. For the same reason $g$ sometimes has three discrete values for the three different directions ($g$ value anisotropy). The $g$ value parallel to the direction of $H_o$ ($g_{||}$) often differs from that in the perpendicular direction ($g_{\perp}$). Both values can be ascertained experimentally.

A second feature of an EPR spectrum is **hyperfine structure** which results from coupling of the magnetic moment of the unpaired electron with nuclear spins. The coupling is analogous to the spin–spin coupling of NMR (Chapter 3). The hyperfine splitting constant A, like the coupling constant $J$ of NMR spectroscopy, is given in Hertz. Splitting may be caused by a magnetic atomic nucleus about which the electron is moving or by some adjacent nucleus or other unpaired electron. Sometimes important chemical conclusions can be drawn from the presence or absence of splitting. Thus, the EPR spectrum of a metal ion in a complex will be split by nuclei in the ligand only if covalent bonding takes place.

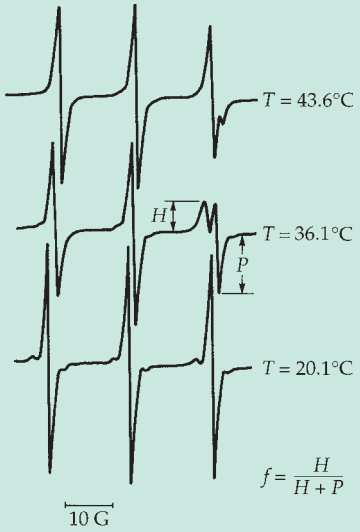
It is customary in EPR spectroscopy to plot the first derivative of the absorption rather than the absorption itself. Thus, for the paramagnetic nitroxide 2,2,6,6-tetramethylpiperidine-1-oxyl the EPR spectrum consists of three equally spaced bands whose peaks are marked at the points where the steep




2,2,6,6-Tetramethylpiperidine-1-oxyl

lines in the middle of the first derivative plots cross the horizontal axis.

Coupling with the $^{14}N$ nuclear spin causes splitting into three lines as shown in the accompanying figure.

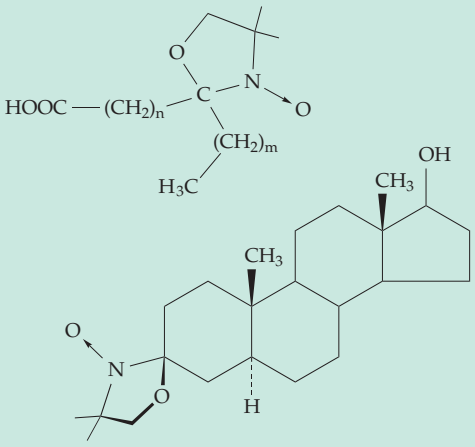



EPR spectrum of tetramethylpiperidine-1-oxyl dissolved in an aqueous dispersion of phospholipids. (Top) above the main bilayer transition temperature $T_t$; (center) between $T_t$ and pretransition temperature; (bottom) below pretransition temperature. From Shimshick and McConnell.[f]

This nitroxide is more soluble in liquid regions of bilayers than it is in solid regions. As bilayers are warmed in the EPR spectrometer, the solubility of this spin-labeled compound in the lipid can be followed (see figure). The lower of the three spectra approximates that of the spin label in water alone, while the others are composite spectra for which part of the spin label has dissolved in the phospholipid bilayers.

**BOX 8-C (continued)**

Since frequencies for EPR spectroscopy are ~100 times higher than those for NMR spectroscopy, correlation times (Chapter 3) must be less than ~$10^{-9}$ s if sharp spectra are to be obtained. Sharp bands may sometimes be obtained for solutions, but samples are often frozen to eliminate molecular motion; spectra are taken at very low temperatures. For spin labels in lipid bilayers, both the bandwidth and shape are sensitively dependent upon molecular motion, which may be either random or restricted. Computer simulations are often used to match observed band shapes under varying conditions with those predicted by theories of motional broadening of lines. Among the many spin-labeled compounds that have been incorporated into lipid bilayers are the following:




Much of the interpretation of the observed changes in EPR spectra of spin labels is empirical. For example, the spectra in the accompanying figure can be interpreted to indicate that the spin label dissolves in the lipid to a greater extent at higher temperatures. The ratio $f$ (defined in the figure) is an empirical quantity whose change can be monitored as a function of temperature. Plots of $f$ vs $T$ have been used to identify transition and pretransition temperatures in bilayers.[f]

EPR spectroscopy is used widely in the study of proteins and of lipid–protein interactions.[c] It has often been used to estimate distances between spin labels and bound paramagnetic metal ions.[g] A high-resolution EPR technique that detects NMR transitions by a simultaneously irradiated EPR transition is known as electron-nuclear double resonance (ENDOR).[h]


[a] Berliner, L. J., and Reuben, J., eds. (1989) *Spinlabeling. Theory and Applications*, Vol. 8, Plenum, New York ((Biological Magnetic Resonance Series)
[b] Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry*, Freeman, San Francisco, California (pp. 525–536, 1352–1362)
[c] Marsh, D. (1983) *Trends Biochem. Sci.* **8**, 330–333
[d] Esmann, M., Hideg, K., and Marsh, D. (1988) *Biochemistry* **27**, 3913–3917
[e] Millhauser, G. L. (1992) *Trends Biochem. Sci.* **17**, 448–452
[f] Shimshick, E. J., and McConnell, H. M. (1973) *Biochemistry* **12**, 2351–2360
[g] Voss, J., Salwinski, L., Kaback, H. R., and Hubbell, W. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12295–12299
[h] Lubitz, W., and Babcock, G. T. (1987) *Trends Biochem. Sci.* **12**, 96–100


fatty acids contain methyl branches (Chapter 21)[147,148] which can decrease the melting point and increase the monolayer surface area by a factor of as much as 1.5. (3) Yet another mechanism for lowering the melting point of fats is the incorporation of **cyclopropane-containing fatty acids** (Chapter 21).

On the other hand, as we have already seen, cholesterol tends to reduce the mobility of molecules in membranes and causes phospholipid molecules to occupy a smaller area than they would otherwise. Myelin is especially rich in long-chain sphingolipids and cholesterol, both of which tend to stabilize artificial bilayers. Within our bodies, the bilayers of myelin tend to be almost solid. Bilayers of some gram-positive bacteria growing at elevated temperatures are stiffened by biosynthesis of bifunctional fatty acids with covalently joined "tails" that link the opposite sides of a bilayer.[149]

Why must membrane lipids be mobile? One reason is probably to be found in the participation of membranes in many vital transport processes. Biological membranes have a relatively high permeability to neutral molecules (including $H_2O$),[64,150] and it has been suggested that above $T_m$ fatty acid chains are free to rotate by 120° around single bonds from trans to gauche conformations. When such rotation occurs about adjacent, or nearly adjacent single bonds, **kinks** are formed. If a kink originates near the bilayer surface, as will usually be the case, a small molecule may jump into the void created. Since the kink can easily migrate through the bilayer, a small molecule may be carried through with it.[151,152] The same factors may assist larger protein molecules which function in membrane transport. They probably also account for the substantial degree of **hydration** of bilayers which involves both the polar head groups and water diffusing through the nonpolar interior.[153]

Not only can molecules diffuse through membranes but also membrane lipids and proteins can move with respect to neighboring molecules. The rates of **lateral**

**diffusion** of lipids in bilayers and of antigenic proteins on cell surfaces are rapid.[80,154] If diffusion of phospholipids is assumed to occur by a pairwise exchange of neighboring molecules, the frequency of such exchanges can be estimated[155] as $\sim 10^7\ s^{-1}$. However proteins may meet many obstacles to free diffusion.[156] Lateral diffusion is often measured by the technique of **fluorescence recovery after photobleaching**. One small spot in a bilayer that contains a dye attached to a lipid or a protein is bleached by a laser beam. Lateral diffusion of nearby unbleached molecules into the bleached spot can then be observed.[80] Lateral diffusion can also be observed by NMR spectroscopy[157] and by single-particle tracking.[158,159] In addition to diffusion there may often be a flow of membrane constituents in directions dictated by metabolism.[160] Although lateral diffusion is fast a "flip-flop" transfer of a phospholipid from one side of the bilayer to the other may require many seconds.[161] However, a sudden increase in the calcium ion concentration, an important intramolecular signal (Chapter 11), activates a "scramblase" protein which promotes a rapid transbilayer movement of phospholipids.[162]

***Electrical properties of membranes.*** Biological membranes serve as barriers to the passage of ions and polar molecules, a fact that is reflected in their high electrical resistance and capacitance. The electrical resistance is usually $10^3$ ohms $cm^{-2}$, while the capacitance is 0.5–1.5 microfarad (μF) $cm^{-2}$. The corresponding values for artificial membranes are $\sim 10^7$ ohms $cm^{-2}$ and 0.6 – 0.9 μF $cm^{-2}$. The lower resistance of biological membranes must result from the presence of proteins and other ion-carrying substances or of pores in the membranes. The capacitance values for the two types of membrane are very close to those expected for a bilayer with a thickness of ~2.5 nm and a dielectric constant of 2.[54,84,163] The electrical potential gradient is steep.

Outer cell surfaces usually carry a net negative charge, the result of phosphate groups of phospholipids, of carboxylate groups on proteins, and of sialic acids attached to glycoproteins. This negatively charged surface layer attracts ions of the opposite charge (counterions), including protons, and repels those of the same charge. The result is development of a diffuse **electrical double layer** consisting of the fixed negative charges on the surface and a positive **ionic atmosphere** extending into the solution for a distance that depends upon the ionic strength.[164–166] This ionic atmosphere is analogous to that postulated by the Debye–Hückel theory (Chapter 6). At the physiological ionic strength of 0.145 M the thickness of the double layer, taken as the distance at which the electrical potential falls to a certain fraction of that at the cell surface,[164,165] is about 0.8 nm. However, the double-layer thickness increases to about three times this value at an ionic strength of $10^{-3}$ M and to still greater distances at lower ionic strengths.

The net surface charge of a cell and the associated electrical double layer are important in interactions between cells and may influence the development of extracellular structure such as basement membranes. The net negative charge on cells also gives rise to an experimentally measurable electrophoretic mobility.

A characteristic of living cells is the maintenance of **ionic gradients** across the plasma membrane. Thus almost all cells accumulate $K^+$, even "pumping" it from very dilute external solutions. Cells also exclude sodium, pumping it out from the cytoplasm by mechanisms considered in Section C,2. If a microelectrode is inserted through a cell membrane and the potential difference is measured between the inside and outside of the cell, a **resting potential** which, in nerve cells, may be as high as 90 mV is observed. The origin of the potential appears to lie in the concentration differences of ions. From the value of $\Delta G$ for dilution of an ion (Eq. 6-25) and the relationship between $\Delta G$ and electrode potential (Eq. 6-63), the Nernst equation (Eq. 8-2) can be derived. According to this equation, which applies to a single ion for which the membrane is permeable,

$$E_m = \frac{RT}{nF} \ln\left(\frac{c_1}{c_2}\right) = \frac{0.059}{n} \log\left(\frac{c_1}{c_2}\right) \quad \text{at } 25°C \tag{8-2}$$

a 10-fold concentration difference across the membrane for a monovalent ion ($n = 1$) would lead to a 59-mV membrane potential, $E_m$. Since membranes are relatively impermeable to sodium ions, it is generally conceded that for many membranes the origin of the membrane potential lies mainly with the potassium ion concentration difference which is maintained by the $Na^+$, $K^+$-ATPase (Section C). A more complete equation takes account of $K^+$, $Na^+$, and $Cl^-$ together with their respective permeabilities.[167–169] Note also that Eq. 6-64 is also often called the Nernst equation.[170]

Protons are also pumped across cytoplasmic and inner mitochondrial membranes, a topic of Chapter 18. The flow of protons from inside to outside also contributes to the membrane potential. The positive charges of $H^+$, $K^+$, and other cations associated with the external membrane surface are balanced by the negative charges of protein molecules as well as $Cl^-$ and phosphate anions that are in or near to the inner surface of the membranes.

Another possibility for proton flow has intrigued biophysicists for years. Membranes often display a substantial electrical conductivity in a lateral direction along the membrane surface.[171–173] Electrical conduction may involve movement of protons along hydrogen-bonded lines, e.g., involving ethanolamine head groups or phosphate groups and bridging water as

previously discussed (see also Eq. 9-96). Alternatively, conduction may depend upon membrane-associated proteins.[174] This lateral proton conduction may be important to many proton-driven membrane processess, such as rotation of bacterial flagella, ATP synthesis, and pumping of ions (Chapter 18).

***The two sides of a membrane.*** Many observations indicate great differences between the inside and outside of the membranes that surround cells.[51,175,176] Bretscher and Raff[62] observed that, among the phospholipids of the erythrocyte membrane, phosphatidylcholine predominates in many mammals but is replaced by sphingomyelin in ruminants. Sheep erythrocytes are resistant to cobra venom phospholipase A, which is known to remove the fatty acid from the central position on the glycerol of phosphatidylcholine, causing lysis of the cells. The resistance of sheep erythrocytes suggested that the sphingomyelin is on the outside of the membrane while the phosphatidylethanolamine and other phospholipids are inside. By inference, phosphatidylcholine is also largely on the outside of plasma membranes. Supporting this conclusion is the observation that most of the reactive amino groups of phosphatidylethanolamine and phosphatidylserine are found on the inside (cytoplasmic) surfaces.[177] Since the total content of phosphatidylcholine and sphingomyelin often exceeds that of phosphatidylethanolamine and phosphatidylserine, the bilayer would be incomplete on the inside of the membrane were it not for the presence of proteins, which contribute more to the inside than to the outside surface.

Glycolipids are usually on the *outside* of plasma membranes with the attached sugar chains projecting into the surrounding water. An important generalization is that *sugar groups attached either to lipids or to proteins tend to be on outer cell surfaces or on materials that are being exported from cells*. An exception is found in the abundant galactolipids of chloroplasts.

## 2. Membrane Proteins

The many proteins present within or attached to membranes have a variety of functions. Some are obviously structural, tying other proteins to a membrane or providing a base for projecting fimbriae, flagella, and other appendages. Some proteins of the outer surface act as anchoring points for macromolecules that lie between cells. The inner surfaces of membranes are attached to the cytoskeleton (Chapter 7). Many of the proteins embedded in membranes control the passage of materials across membranes. Others serve as receptors that sense the presence of specific compounds or of light. Membranes may also contain foreign proteins such as subunits of virus coats. Proteins that are deeply embedded in membranes are referred to as **intrinsic** or **integral membrane proteins**. Proteins that are more loosely associated with the membrane, principally at the inner surface, are called **peripheral**.[61]

***Integral membrane proteins.*** Membrane proteins are hard to crystallize[178] and precise structures are known for only a few of them.[179–181] A large fraction of all of the integral membrane proteins contain one or more **membrane-spanning helices** with loops of peptide chain between them. Folded domains in the cytoplasm or on the external membrane surface may also be present. The best-known structure of a transmembrane protein is that of the 248-residue bacteriorhodopsin. It consists of seven helical segments that span the plasma membrane (Fig. 23-45) and serves as a light-activated proton pump. Other proteins with similar structures act as hormone receptors in eukaryotic membranes. A seven-helix protein embedded in a membrane is depicted in Fig. 8-5 and also, in more detail, in Fig. 11-6.

The most hydrophobic integral membrane proteins can be extracted into organic solvents such as mixtures of chloroform and methanol. One such **proteolipid protein**, the 23.5-kDa lipophilin, accounts for over half the protein of myelin.[57,182] The purified protein from rat brain contains 66% of nonpolar amino acids and six molecules of covalently bound palmitic acid and other fatty acids per peptide chain in thioester linkage to cysteine side chains. This protein evidently has four transmembrane helical segments with the six fatty acid chains incorporated into the membrane bilayer. It also has cytoplasmic and extracellular loops, one of which binds inositol hexakisphosphate (Ins *P*-6). (Fig. 11-9).[183] The myelin proteolipid is an essential component of the myelin sheath and defects in this protein are associated with some demyelinating diseases[57] which are discussed in Chapter 30.

There are many known topologies for helix-bundle membrane proteins with the number of membrane-spanning helices ranging from 1 to 14 or more e.g., see Fig. 8-23.[184–186] A topology can often be predicted using suitable computer programs.[187–191] A first step is to identify all sequences of 20 or more residues that could form a helix sufficiently hydrophobic to allow good nonpolar interactions with the bilayer core. Sequences that can form **amphipathic** (amphiphilic) helices must also be considered because two or more of these can pack together in a membrane with their hydrophilic sides together, sometimes forming pores (see Section C,1).[192]

Predictions of membrane protein structure are in part based on the **positive-inside rule** which states that positively charged lysine and arginine residues will not pass through a membrane but will remain on the negatively charged cytoplasmic surface. Often, N-terminal parts of a protein will pass through a

membrane and will contain glutamate or aspartate residues whose side chains carry negative charges. These tend to remain outside of the membrane where they are attracted to the positive charges on the membrane outer surface. Positively charged residues may interact with phosphate groups of phospholipids and tyrosine and tryptophan side chains may interact with the carbonyl groups of the ester linkages.[192a,b]

After probable transmembrane helical regions have been identified a residue-by-residue attempt can be made to identify cytosolic and extracellular loops. Chemical reactivities of the naturally occurring side chains can also be examined. Residues within the helix bundle will be protected. Mutations can be prepared systematically by "**scanning mutagenesis**." For example, presumed extracellular loops in the erythrocyte band 3 protein (next section) have been mutated by introduction of *N*-glycosylation acceptor sites (Asn-X-Ser / Thr). If the sequence is at least 12–14 residues away from transmembrane sequences it will probably be glycosylated in a suitable laboratory test system.[193] Within a transmembrane domain amino acid residues can be systematically replaced with alanine ("alanine scanning mutagenesis") or with cysteine and effects on the protein can be observed. Substitution with cysteine allows another possibility: Two cysteines on adjacent transmembrane helices may be linked as disulfides if they are close enough.[194,195] Discovery of such neighboring pairs can be very valuable in attempting to establish relationships of one helix to another. A cloned gene can also be split into two pieces prior to crosslinking of cysteine side chains. This may facilitate mapping of tertiary interactions within transmembrane proteins.[196] Molecular dynamics methods for modeling helix bundle proteins are also being developed.[197]

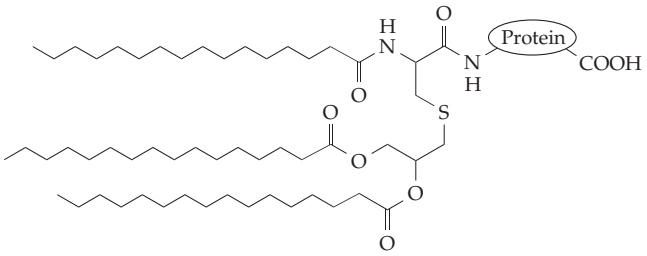
Not all integral membrane proteins have a helix bundle structure. Some of the simplest transmembrane proteins are the subunits of bacterial viruses and pili (Figs. 7-7 and 7-9) The 7-nm α-helical rods of phage M13 have a 20-residue hydrophobic section (residues 25-46) which is preceded by a negatively charged sequence that appears to form a short *amphipathic* helix that lies on the external surface of the membrane and helps to anchor the subunit. A positively charged cluster at the opposite end of the hydrophobic helical region remains in the cytoplasm.[180,186,198] The **porins**, which form pores in outer membranes of bacteria and ribosomes, are large 16-strand β cylinders (Section C,1).[179,180] It has been suggested that this different basic architecture may be related to the fact that the porin polypeptides must be exported through the inner membrane before being refolded in the periplasm.[199] Keep in mind that a β strand of nine residues can span the 33-nm bilayer core just as well as can a 22-residue helix. Beta structures as well as shorter helices may well be present in proteins that also contain membrane-spanning helices.

A third important structural pattern involves extensive use of amphipathic helices that lie partially embedded in a membrane surface. For example, the blood lipoproteins are lipid particles partially coated by amphipathic helices (Chapter 21).[200,201]

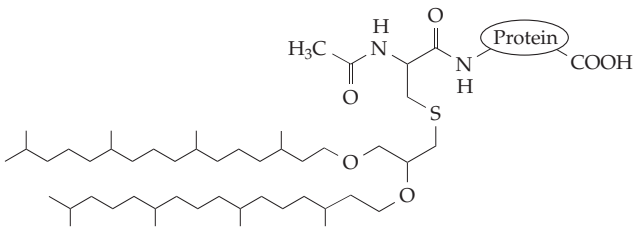
***Anchors for proteins.*** Proteins with membrane-spanning sequences are usually anchored into the membranes with the help of many polar side chains, often including glycosylated residues, in the ends and loops that protrude on the two sides of the membranes. Other proteins are anchored by insertion into membranes of nonpolar groups. These include fatty acyl groups at N termini, fatty acid ester groups on serine, threonine, or cysteine side chains, and polyprenyl groups attached to cysteine side chains as thioethers.[202] Many membrane-anchored proteins carry the saturated 14-carbon **myristoyl** group in amide linkage with an N-terminal glycine of the protein.[203–206] The fatty acid chain is added at the time of the protein synthesis, i.e., **cotranslationally**. It can intercalate into the membrane bilayer but provides a relatively weak anchor.[207] The 16-carbon **palmitoyl** group is typically added to a cysteine thiol in recognition sequences at various positions in a protein.[182,202,208,209] The modification occurs after protein synthesis, i.e., **posttranslationally**. Because thioesters are relatively unstable, palmitoylation is regarded as a reversible modification.[210]

A third lipid anchor is provided by the polyprenyl **farnesyl** (15-carbon) and **geranylgeranyl** (20-carbon) groups in thioether linkage to cysteine residues. These must be present in specific recognition sequences at the C termini of proteins, most often with the sequence CAAX.[211–215] The prenylation (also called isoprenylation) reaction is followed by proteolytic removal of the last three residues (AAX) and methylation of the new C-terminal carboxyl group as is discussed in Chapter 11, Section D,3. See also Chapter 22, Section A,4.

A lipoprotein present in the periplasmic space of *E. coli* is anchored to the outer bacterial membrane by a triacylated modified N-terminal cysteine containing a glyceryl group in thioether linkage as shown in the following structure (see also Section E,1).




A related anchor that uses diphytanylglycerylation is found in certain proteins of archaeobacteria.[216]




A series of **glycosylphosphatidylinositol** (GPI) anchors that are covalently linked to a variety of proteins utilize diacylglycerol or alkylacylglycerol for attachment to a bilayer. The proteins are joined through their C-terminal carboxyl groups to the diacylglycerol by a chain of covalently linked ethanolamine, phosphate, mannose, glucosamine, and *myo*-inositol as shown in Fig. 8-13.[217–223] The proteins are linked to the diacylglycerol through the conserved structure: $H_2N$-protein→ethanolamine→*P*→6Manα1→2Manα1→6Manα1→4GlcNAc→Ins→diacylglycerol.

The structure in Fig. 8-13, which anchors the small Thy-1 antigen to surfaces of rat thymocytes,[224] contains additional mannose, *N*-acetylgalactosamine, and ethanolamine phosphate. These groups may be missing or substituted by other groups in other anchors.[225] For example, that of human erythrocyte acetylcholinesterase lacks the extra mannose and GalNAc of the Thy-1 anchor but contains a palmitoyl group attached to an oxygen atom of the inositol. This provides an additional hydrocarbon tail that can enter the bilayer.[218] Other PI-anchored proteins include enzymes, such as alkaline phosphatase and lipoprotein lipase, and surface proteins of the parasites *Trypanosoma*,[225]*Leishmania*, and *Toxoplasma*.[220] Some adhesion molecules and a variety of other outer surface proteins are similarly attached to membranes. Analysis of the genome of *Caenorhabditis elegans* suggests that the nematode contains over 40 GPI-tailed proteins and perhaps more than 120.[225a]

***Analyzing erythrocyte membranes.*** The proteins of red blood cell membranes were among the first to be studied. Because membrane proteins are present in small amounts and tend to be hard to dissolve without denaturation, they have been difficult to study. They usually do not dissolve readily in water, but red cell membranes can be almost completely solubilized in water using a $5 \times 10^{-3}$ M solution of the chelating agent EDTA (Table 6-9) or by 0.1 M tetramethyl ammonium bromide.[226] These observations suggested that ionic linkages between proteins, or between proteins and phospholipids, are important to membrane stability. Nonionic detergents such as those of the Triton X series or β-octylglucopyranoside also solubilize most membrane proteins,[178,227–229] whereas ionic detergents such as sodium dodecyl sulfate (SDS) often cause unfolding and denaturation of peptide chains.

Gel electrophoresis of plasma membrane proteins in SDS solution yields ~10 prominent bands and at

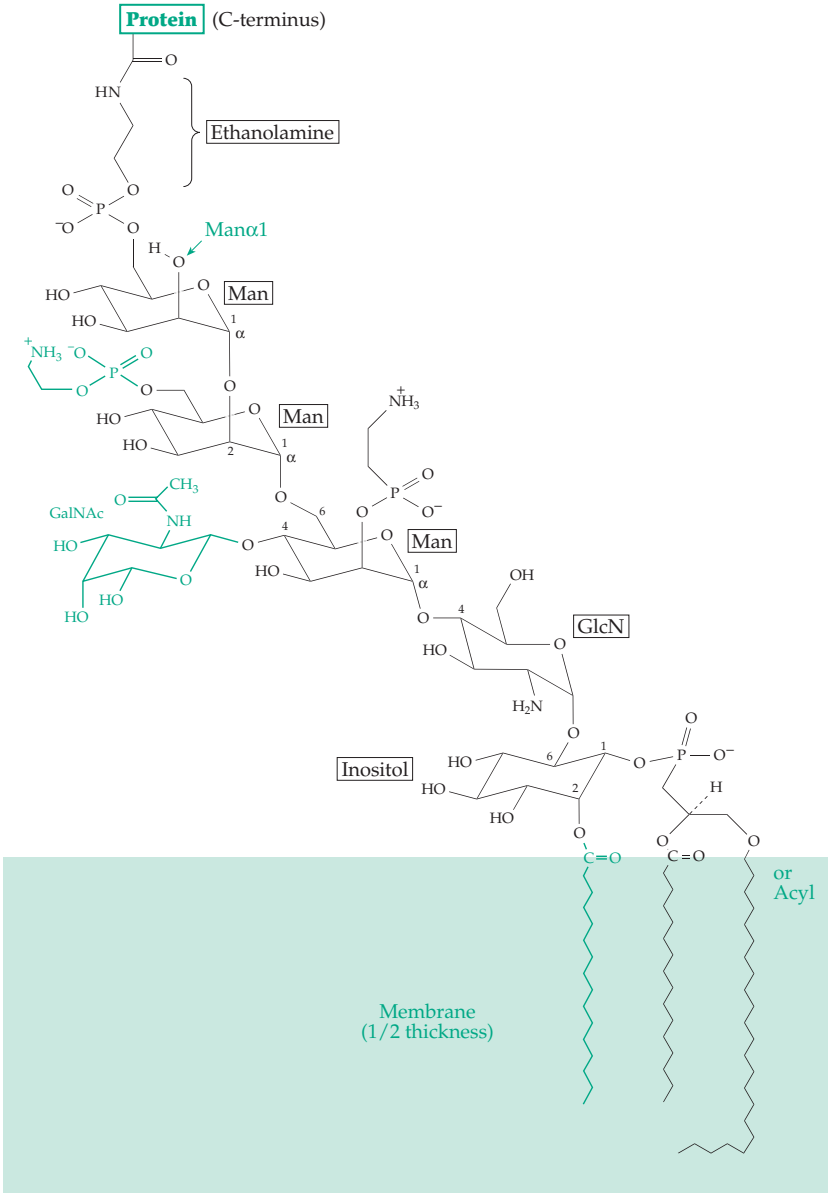
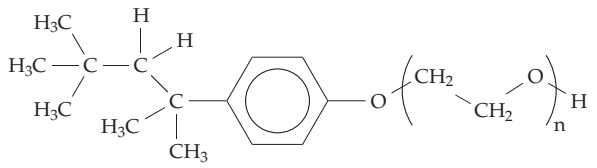



**Figure 8-13** Structure of glycosylphosphatidylinositol (also called phosphatidylinositol-glycan) membrane anchors. The core structure is shown in black. The green parts are found in the Thy-1 protein and / or in other anchors.



Polyoxyethylene p-t-octyl phenol: the Triton X series of detergents. n = 9–10 for Triton X-100 and 7–8 for Triton X-114.

least 30 less intense bands ranging in molecular mass from 10 to 360 kDa.[230,231] Proteins of erythrocyte ghosts, and the usual system for numbering them, are shown in Fig. 8-14. Some very important proteins known to be present in this membrane, such as $(Na^+ + K^+)$-activated ATPase (Section C,2), are found in such low quantities (e.g., a few hundred molecules in a single red blood cell)[62] that they do not show up in electropherograms. Mitochondrial membranes appear to be more complex than plasma membranes, but myelin has a somewhat simpler composition.

***Glycoproteins.*** Many of the integral proteins of membranes are glycoproteins.[234–236] These may sometimes be recognized in electropherograms because they are stained by the periodic acid-Schiff (PAS) procedure. At least 20 glycoproteins are present in erythrocyte ghosts, and glycoproteins appear to be prominent protein components of the plasma membranes of all eukaryotes and of primitive archaebacteria such as *Halobacterium*.

The most abundant glycoprotein of red blood cell membranes is the 95-kDa PAS-reactive **band 3 protein** (Fig. 8-14) which makes up ~25% of the total membrane protein.[236–240] A variety of asparagine-linked oligosaccharides based on the core hexasaccharide structure shown in Chapter 4, Section D,2 are present and apparently project into the surrounding medium. Electron micrographs of freeze-fractured surfaces through the membrane bilayer (Fig. 8-15) show ~4200 particles of 8 nm diameter per square micrometer, randomly distributed and apparently embedded in the membrane. These probably represent dimers of the glycoprotein. The amino acid sequence of the 911-residue protein suggests 13 membrane-spanning helices in the 550-residue C-terminal domain and that the 41-kDa N-terminal domain projects inward into the cytoplasm.[236,240] Among the first 31 residues are 16 of aspartate or glutamate. These provide a highly negatively charged tail that is able to interact electrostatically with other proteins.[236] Among these are components of the cytoskeleton, which appears to be anchored to the membrane via the band 3 protein. The band 3 protein is also a substrate for transglutaminase (Eq. 2-23) which creates covalent crosslinks to other proteins.[237] Another major function of the band 3 glycoprotein is to form channels for the transport of anions (Section C,2).

Another integral glycoprotein of erythrocytes, the 31-kDa **glycophorin A** (PAS-1),[235,241–244] is ~60% by

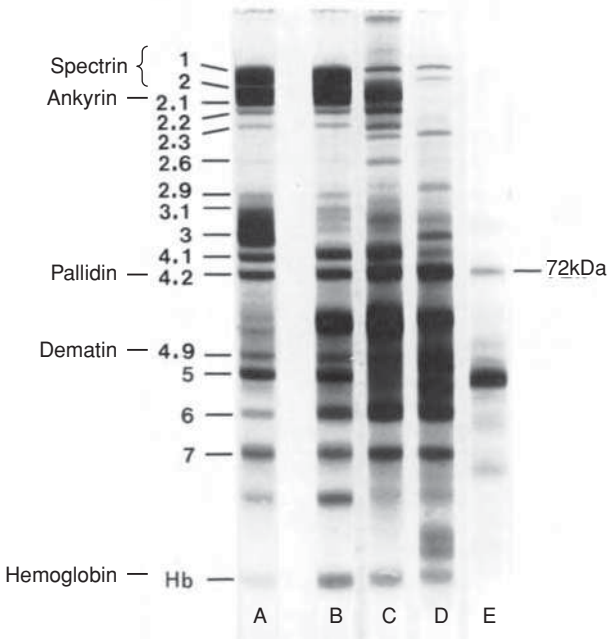



**Figure 8-14** SDS-polyacrylamide gel electrophoresis of human erythrocyte ghosts. (A) From untreated cells. (B) From cells digested externally with chymotrypsin. (C) Inside-out vesicles prepared from cells pretreated with chymotrypsin. (D) The same inside-out vesicles after further treatment with chymotrypsin. (E) Polypeptides released by the chymotryptic treatment of the inside-out vesicles. The peptides are numbered according to the system of Steck[232]; Hb, hemoglobin. From Luna *et al*.[233]

weight carbohydrate. Its 131-residue chain has a single ~23-residue membrane-spanning helix. The first 50 residues at the N terminus, which project from the outside surface of the membrane, include many serines and threonines. Their side chains carry 15 O-linked tetrasaccharides and one complex N-linked oligosaccharide. There are a total of ~ 160 sugar residues per peptide chain, largely *N*-acetylgalactosamine, galactose, and sialic acid. Some of the oligosaccharides contain MNO blood group determinants[234,241] (Box 4-C). Because of a high content of sialic acid, glycophorin also carries a large negative charge. The 35-residue C-terminal domain is hydrophilic and rich in proline, glutamate, and aspartate. It probably extends into the cytoplasm and may bind calcium ions or interact with $-NH_3^+$ groups on phospholipid heads.

If all the sugar residues of the glycophorin molecules in an erythrocyte were spread over the surface of the cell they could cover approximately one-fifth of its surface in a loose network. However, it is more likely that they form bushy projections of a more localized sort. These oligosaccharides not only act as immunological determinants but also serve as receptors for influenza viruses. Other glycoproteins related to glycophorin A occur in smaller amounts.[244]

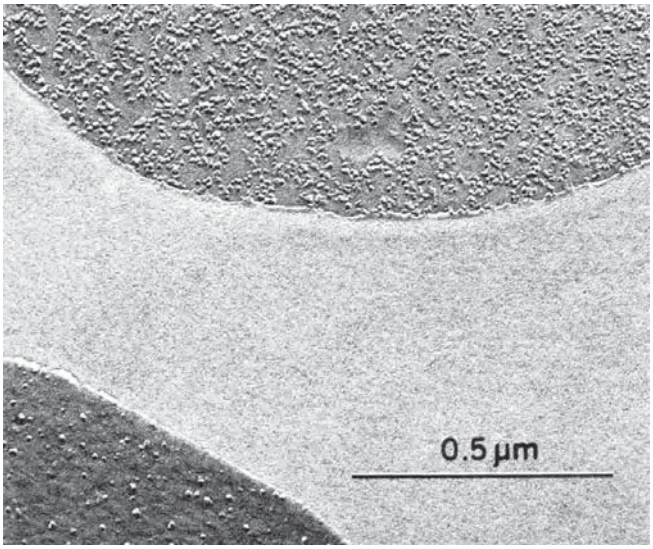



**Figure 8-15** Freeze-fractured membranes of two erythrocyte "ghosts." The upper fracture face (PF) shows the interior of the membrane "half" closest to the cytoplasm. The smooth region is lipid and contains numerous particles. The lower face, the extracellular half (EF), possesses fewer particles. The space between the two is nonetched ice. See Figs. 1-4 and 1-15 for electron micrographs of sections through biological membranes. Courtesy of Knute A. Fisher.

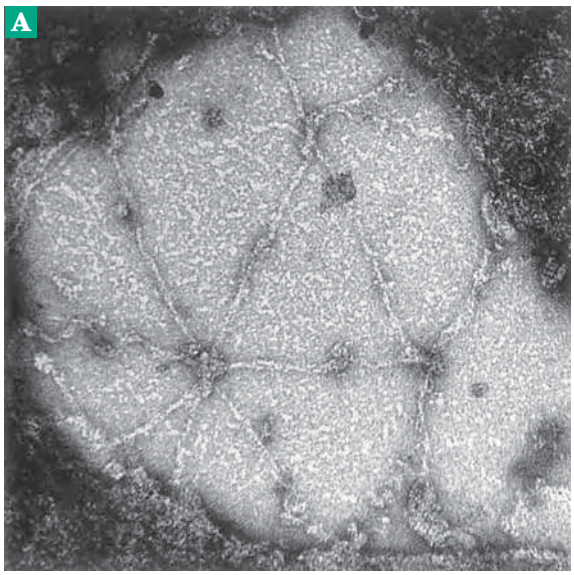
***Connections to the cytoskeleton.*** About one-third of the protein of the red blood cell membrane is accounted for by a pair of larger hydrophobic peptides called **spectrin** with molecular masses of 280 kDa (α chain) and 246 kDa (β chain).[245–250a] These are found in bands 1 and 2 of Fig. 8-14. The spectrin monomers consist largely of 106- to 119- residue repeat sequences each of which folds into a short triple-helical bundle (Fig. 8-16B). The beaded-chain monomers associate readily to αβ dimers, long ~100 nm thin flexible rods which associate further to $(\alpha\beta)_2$ tetramers. The latter, in turn, bind to monomers or to small oligomers of actin. In red blood cells the actin crosslinks the spectrin tetramers into a two-dimensional "fishnet" (Fig. 8-16A), the ~85,000 spectrin tetramers uniformly covering the entire inner surface (130 $\mu m^2$) of the erythrocyte.[245,251] The inner location of spectrin was established by the fact that chemical treatments[252] that covalently label groups on proteins exposed on the outer surface of erythrocytes (e.g., iodination with lactoperoxidase; Chapter 16) did not label spectrin.[257]

Spectrin also binds to one domain of another large 215-kDa peripheral protein called **ankyrin** (band 2.1, in Fig. 8-14) which anchors the spectrin network to the membrane.[258,259] Ankyrin is actually a multigene family of related proteins that are present in many metazoan tissues.[258,260] These are modular proteins with separate binding domains for spectrin and band 3 protein. The latter domain contains 24 **ankyrin repeats**, 33-residue modules, also found in a variety of other proteins.[258,259]

Ankyrin binds firmly to the band 3 glycoprotein which is embedded in the membrane. The 78 kDa band 4.1 protein is another major component of the erythrocyte **membrane skeleton**.[261,262] Protein 4.1 binds both to ankyrin and also to **glycophorin C**, providing another anchor to an integral membrane protein (Fig. 8-16C). Spectrin:actin:protein 4.1 in a 1:2:1 ratio are the major components of the membrane skeleton.[262] Other less abundant proteins include **adducin**, protein 4.2 (**pallidin**, which interacts with band 3 protein),[263] protein 4.9 (**dematin**, an actin bundling protein),[264] and the muscle proteins tropomyosin and tropomodulin (Chapter 19). While spectrin and ankyrin of erythrocyte membranes have been studied most intensively, related proteins occur in other cells.[265] Spectrin of brain and other tissues is also known as **fodrin**.[266] **Dystrophin** and **α-actinin**, actin-crosslinking proteins of muscle, are also members of the spectrin superfamily.[248] Protein 4.1 also occurs in various organisms,[267] in various tissues, and in various locations in cells.[268]

What is the function of the membrane skeleton? There is a group of hereditary diseases including **spherocytosis** in which erythrocytes do not maintain their biconcave disc shape but become spherical or have other abnormal shapes and are extremely fragile.[269–272] Causes of spherocytosis include defective formation of spectrin tetramers and defective association of spectrin with ankyrin or the band 4.1 protein.[265,273] Thus, the principal functions of these proteins in erythrocytes may be to strengthen the membrane and to preserve the characteristic shape of erythrocytes during their 120-day lifetime in the bloodstream. In other cells the spectrins are able to interact with microtubules, which are absent from erythrocytes, and to microtubule-associated proteins of the cytoskeleton (Chapter 7, Section F).[270] In nerve terminals a protein similar to erythrocyte protein 4.1 may be involved in transmitter release.[274] The cytoskeleton is also actively involved in transmembrane signaling.

***Integrins and focal adhesions.*** Mature erythrocytes have no nucleus and lack the microtubules and actin filaments that span other cells. In nonerythroid cells the major connections of the cytoskeleton to the membrane are through large αβ heterodimeric membrane-spanning proteins called integrins (Fig. 8-17). These proteins, as well as the ends of the actin filaments, tend to be concentrated in regions long observed and described by microscopists as focal adhesions.[275–279] These are also sites of interaction with the external proteins that form the **extracellular matrix (ECM)**. There are at least 16 different integrin



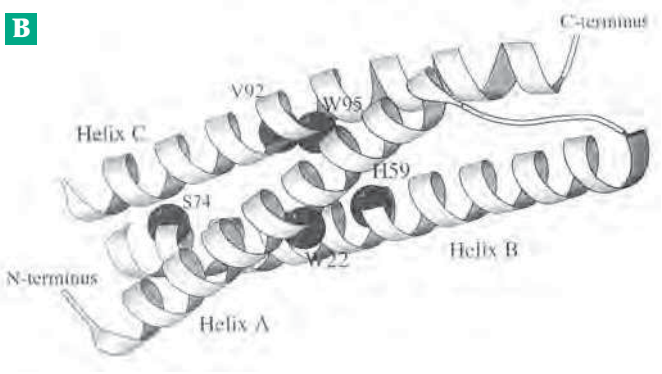
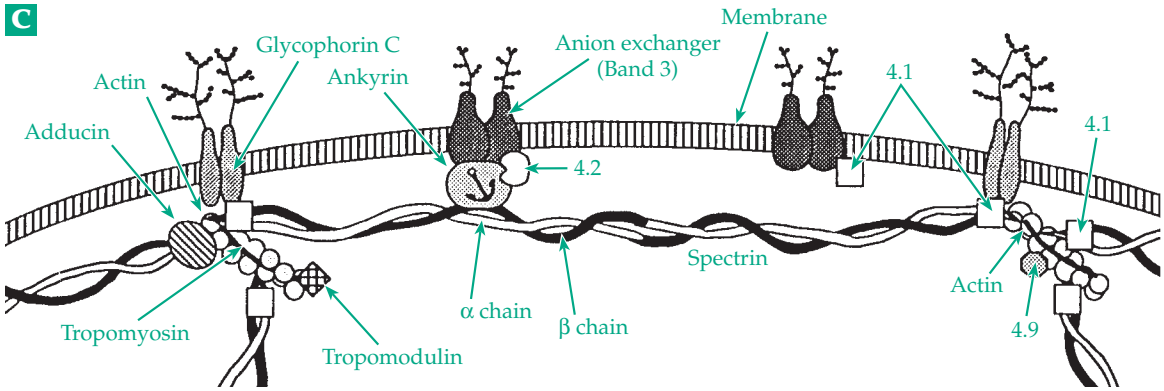




**Figure 8-16** The erythrocyte membrane skeleton. (A) Electron micrograph showing a region of the membrane skeleton (negatively stained, X 200,000) and artificially spread to a surface area nine to ten times as great as in the native membrane. Spreading makes it possible to obtain clear images of the skeleton whose protein components are so densely packed and so subject to thermal flexing on the native, unspread membrane that it is difficult to visualize the individual molecules and the remarkably regular way that they are connected. The predominantly hexagonal and pentagonal network is composed of spectrin tetramers cross-linked by junctions containing actin oligomers and band 4.1 protein. Band 4.9 protein and tropomyosin are probably also bound to the oligomers, whose length (13 actin monomers long) corresponds to the length of a tropomyosin molecule. From Byers and Branton.[253,254] Courtesy of Daniel Branton. (B) Proposed triple α-helical structure of a single spectrin repeat unit.[255] Courtesy of Ruby I. MacDonald and Alfonso Mondragón. Each α spectrin chain consists of 20 and each β spectrin chain of 17 such repeats, which have only partially conserved sequences. The α and β chains are thought to associate in a side-by-side fashion and the αβ heterodimers in an end-to-end fashion to give tetramers. These are the rod-like structures seen in (A) and (C). (C) Cross section of the unspread membrane and cytoskeleton as pictured by Luna and Hitt.[256] Major interactions among components of the cytoskeleton are shown. Apparent sizes of the protein subunits, based on migration positions in SDS-polyacrylamide gel electrophoresis are: spectrin (260 and 225 kDa), adducin (105 and 100 kDa), band 3 (90 to 100 kDa), protein 4.1 (78 kDa), protein 4.2 (pallidin, 72 kDa), dematin (protein 4.9, 48 kDa), glycophorin C (~25 kDa), actin (43 kDa), and tropomyosin (29 and 27 kDa). Courtesy of Elizabeth J. Luna.




α subunits, whose external domains consist of up to 1114 residues, at least 8 β subunits with external domains of up to 678 residues, and at least 22 distinctly different αβ heterodimers.[280–282] The N-terminal part of each α subunit contains seven repeats of ~60 residues each, probably arranged as a β-propeller (see Fig. 11-7D or 15-23).[281] The integrins are structurally complex. Some contain a nucleotide-binding domain (see Figs. 2-13 and 2-27C).[283] Integrins have differing and quite exacting specificities toward the proteins of the external matrix to which they bind. They span the cell membrane and appear to be actively involved in communication between the cytoskeleton and external proteins.[275,277,284] They are often described as **receptors** for the proteins that bind to them.

***Other cell adhesion molecules.*** Long before the discovery of integrins another class of transmembrane adhesion molecules were recognized as members of the immunoglobulin family. These were called cell

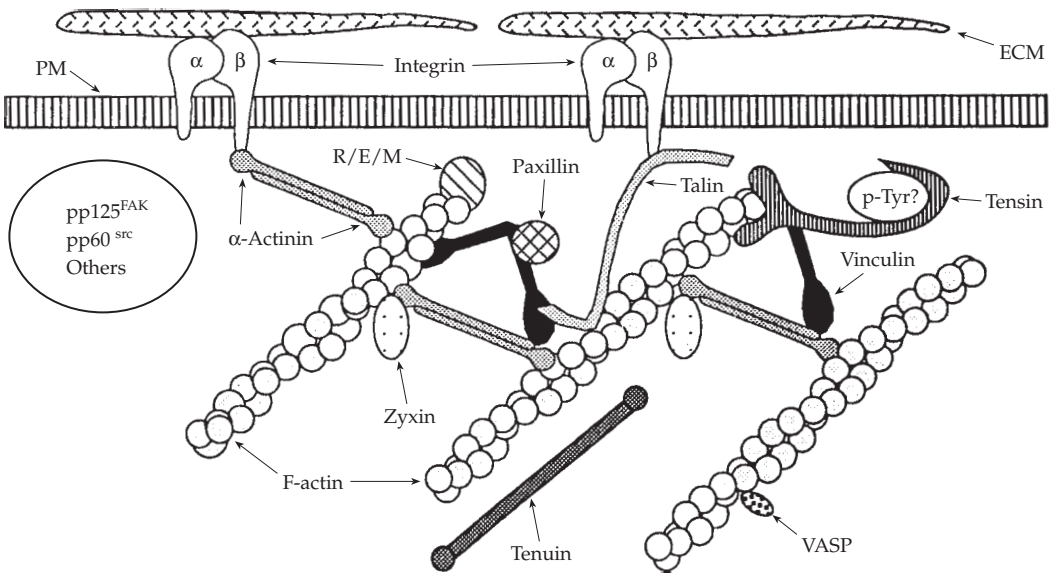



**Figure 8-17** Working model of the protein–protein interactions in focal adhesions determined by *in vitro* binding experiments and immunolocalization. In addition, several interactions are of relatively low affinity in solution but may be enhanced at the membrane surface. Abbreviations are: ECM, extracellular matrix; PM, plasma membrane; p-Tyr-?, unknown phosphotyrosine-containing protein; R/E/M, member of the radixin/ezrin/moesin family; VASP, vasodilator-stimulated phosphoprotein. Diagram is modified from Simon *et al.*[285]

adhesion molecules (**CAMs**) by Edelman.[286,287] These molecules go by names such as intercellular adhesion molecule-1 (ICAM-1), neural cell adhesion molecule (NCAM),[288,289] liver cell adhesion molecule (LCAM), vascular cell adhesion molecule (VCAM), and platelet endothelial cell adhesion molecule (PECAM).[290,291] Many of these proteins, including the **T-cell antigen CD2**,[291a] were first recognized on leukocyte surfaces as differentiation antigens and are often designated by the "Cluster of Differentiation" names (Chapter 31). For example, ICAM-1 is also called CD54 and PECAM CD31.[290,292] The genes for these proteins are often expressed differentially in various tissues and the mRNA molecules formed undergo alternative splicing.[287,289] The extracellular domains of these adhesion molecules consist largely of Ig domains and most have a single transmembrane helix and a small cytoplasmic C-terminal domain. Some of the alternatively spliced forms are attached to the membrane by PGI tails.[287,289] ICAM-1 (Fig. 8-18) has five Ig domains and VCAM-1 has seven,[293] but some CAMs have only two.[294] The CAMs are glycoproteins, often with large *N*-linked oligosaccharides attached. The widely distributed N-CAM contains long α2→8 linked polysialic acid chains on two of the three *N*-glycosylation sites in the fifth Ig domain.[289] The CAMs are often referred to as **receptors**. Their ligands include surface proteins such as fibronectin (next section) but also the integrins, which are also called receptors. Integrins are **coreceptors** for receptors of the Ig superfamily. Each of the coreceptors in a pair has binding sites for other ligands as well (Fig. 8-18). The CAMs and many other adhesion molecules are most abundantly expressed in embryonic tissues in which cells often migrate to new locations and for which the communication with neighboring cells is especially active. Many adhesion molecules bind only weakly and reversibly to their ligands, allowing cells to move.

The **cadherins** are calcium-dependent adhesion proteins that mediate direct cell–cell interactions.[295,296] The external parts of the cadherins also have repeated structural domains with the Ig fold.[297-298b] They have high affinity for each other, allowing cadherins from two different cells to interact and tie the cells together with a zipper-like interaction that is stabilized by the bound $Ca^{2+}$ ions,[297,300] and may be relatively long-lived. The gene for cadherin E is often mutated in breast cancers and may be an important **tumor suppressor gene** (Box 11-D).[301]

***Peripheral proteins of the outer membrane surface.*** Many integral membrane glycoproteins have their sugar-bearing portions exposed on the outer surface of the plasma membrane. Among these are receptors, ion pumps, and biochemical markers of individuality. In addition to these proteins, which are actually embedded in the bilayer, there are external peripheral proteins. One of the best known of these is

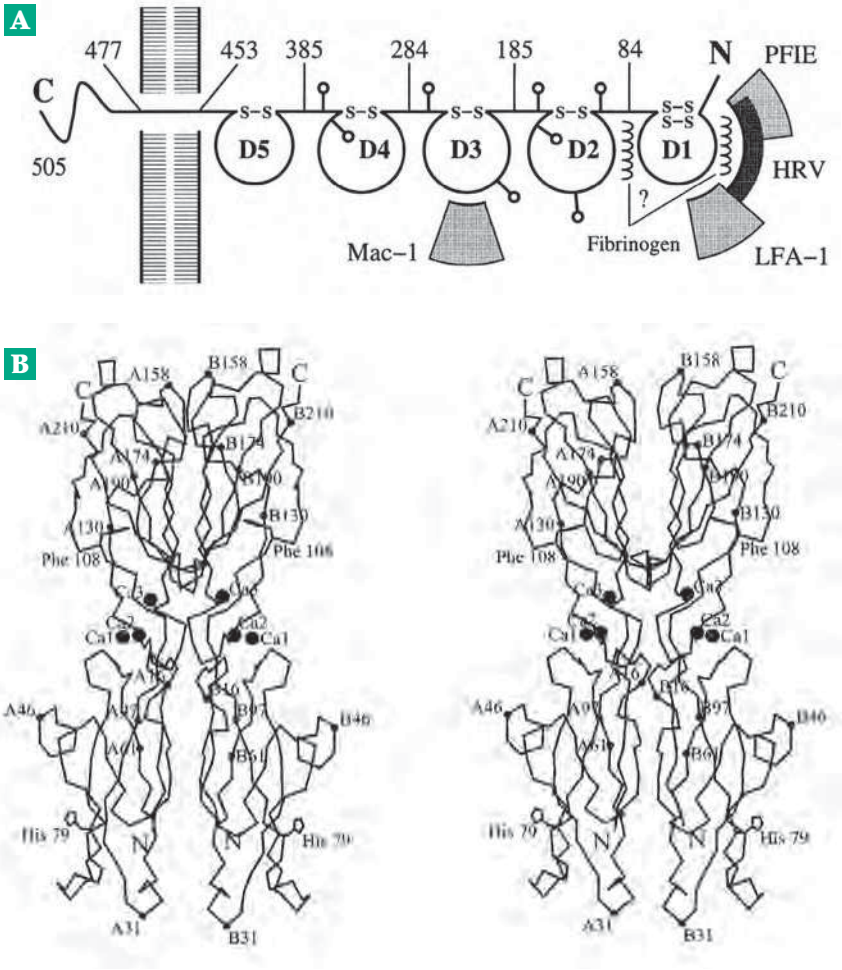



**Figure 8-18** (A) Diagram of the ICAM-1 molecule. The structures labeled D1–D5 are the Ig domains. The glycosylation sites are labeled with small lollipops and approximate sites for binding of the α chain (CD11a) of the integrin called LFA-1, for the macrophage antigen Mac-1, for fibrinogen, for a ligand from erythrocytes infected by the malaria organism *Plasmodium falciparum* (PFIE), and for human rhinoviruses (HRV) are labeled. The binding sites are indicated schematically but each one is a complex interacting surface complementary to its ligand. From Bella *et al.*[299] Courtesy of Michael Rossman. (B) Structure of two N-terminal domains of E-cadherin. The molecules of the dimer are related by a noncrystallographic twofold symmetry axis running vertically in the plane of the page. Clusters of three calcium ions are bound in the linker regions, connecting the N- and C-terminal domains of each molecule and, in this view, are separated by the twofold axis. The N- and C-terminal domains are composed of seven-stranded β-barrels showing the same topology and similar three-dimensional structures. From Nagar *et al.*[296]

**fibronectin** (from *fibra*, "fiber", and *nectare*, "to bind").[302–309] This very large 470-kDa glycoprotein is a disulfide-linked dimer. Appearing under the electron microscope as having two 60-nm arms,[305] fibronectin molecules join together and surround animal cells, anchoring other proteins and carbohydrates of the ECM to the cells. Fibronectin binds tightly to several different cell surface integrins and also to collagen and to glycosaminoglycans of the matrix. It binds to the blood-clotting protein fibrinogen, to actin, and also to staphylococci and other bacteria. Each peptide chain of the fibronectin dimer is organized as several domains. Fibrin and staphylococci bind to the N-terminal domain. A second domain binds collagen and a domain near the C-terminus binds heparin. Several different integrins of cell surfaces bind to the region of the 8th, 9th, and 10th type III repeats.[308] The specificity of the binding depends to a large extent on the presence of the specific tripeptide sequence Arg-Gly-Asp (**RGD**) in a type III repeat (Fig. 8-19B). Peptide sequences such as the PHSRN (Fig. 8-19B) and others[310] also participate in binding to specific integrins. Initial interactions are noncovalent but fibrin and collagen gradually become covalently attached to fibronectin through isopeptide linkages formed by transglutaminase (Eq. 2-23). This enzyme is abundant

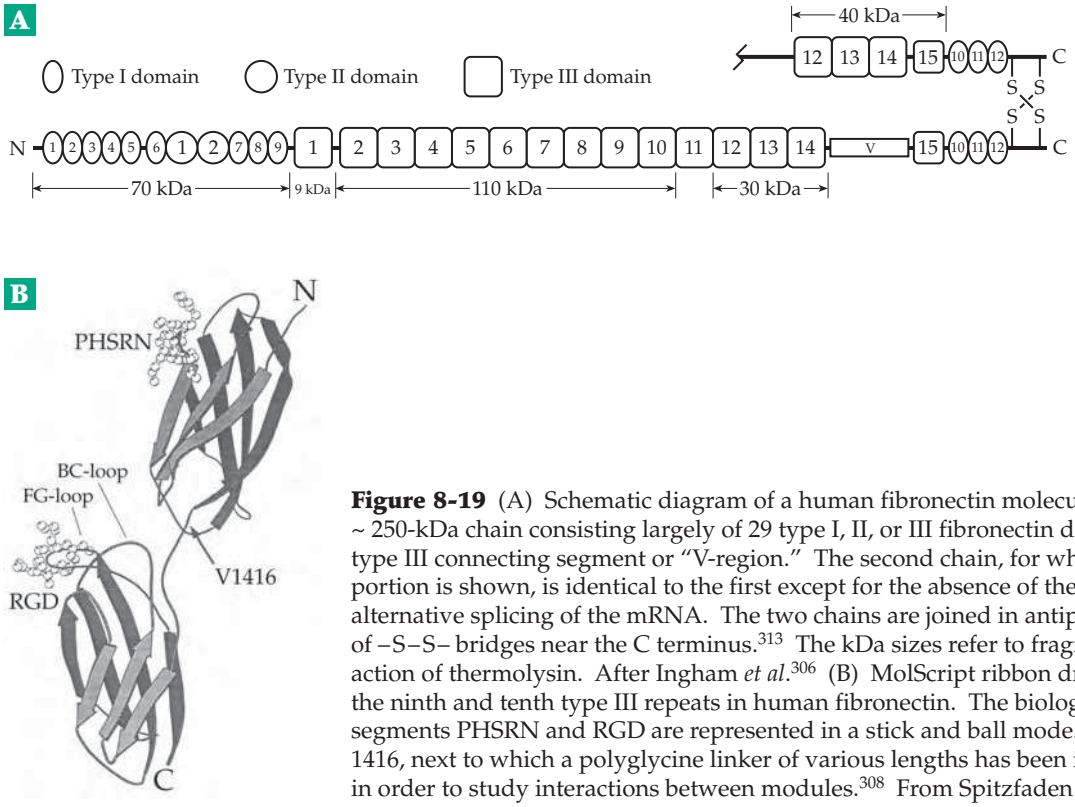



**Figure 8-19** (A) Schematic diagram of a human fibronectin molecule showing one complete ~ 250-kDa chain consisting largely of 29 type I, II, or III fibronectin domains with a 12-kDa type III connecting segment or "V-region." The second chain, for which only the C-terminal portion is shown, is identical to the first except for the absence of the V-region as a result of alternative splicing of the mRNA. The two chains are joined in antiparallel fashion by a pair of –S–S– bridges near the C terminus.[313] The kDa sizes refer to fragments generated by the action of thermolysin. After Ingham *et al.*[306] (B) MolScript ribbon drawing of the structure of the ninth and tenth type III repeats in human fibronectin. The biologically active peptide segments PHSRN and RGD are represented in a stick and ball mode. The arrow marks valine 1416, next to which a polyglycine linker of various lengths has been inserted experimentally in order to study interactions between modules.[308] From Spitzfaden *et al.*[314]

in many developing tissues but often disappears as the tissues mature.[311,312] Fibronectin was formerly known as the "large external transformation-sensitive protein," a name derived from the fact that its quantity is greatly reduced in many virus-transformed cells. This might explain the loss of adhesiveness and of "contact inhibition" observed for cancer cells (Box 11-D).

There are 20 isoforms of human fibronectin. These arise by alternative splicing of the primary gene transcript. Their formation is developmentally regulated.[308,315] One of the isoforms is present as a soluble protein (cold-insoluble globulin) in blood plasma.[302] The fibronectins belong to a larger family of **cytoadhesins**, among which are the blood-clotting proteins **fibrinogen** and **von Willebrand factor** (Chapter 12), and also **thrombospondin**, **vitronectin**,[316] **tenascin**,[317–320] **laminin**,[321,322] **osteopontin**,[323] and **collagens**. These proteins all have a modular construction with repeated domains, often of several different types.[324] For example, tenascin contains 14 ½ EGF domains (Table 7-3), 16 fibronectin type III repeats, and a C-terminal segment homologous to fibrinogens.[325] Laminin (see Fig. 8-33) also has EGF-like modules. Most of these proteins also have in common the presence of the RGD sequence, which binds the cytoadhesins to those integrins and other molecules that carry RGD receptor sites.[304,326–328] The extracellular matrix is discussed further in Section E,2.

## 3. Enzymes and Membrane Metabolism

Many of the proteins of membranes are enzymes. For example, the entire electron transport system of mitochondria (Chapter 18) is embedded in membranes and a number of highly lipid-soluble enzymes have been isolated. Examples are **phosphatidylserine decarboxylase**, which converts phosphatidylserine to phosphatidylethanolamine in biosynthesis of the latter, and **isoprenoid alcohol phosphokinase**, which participates in bacterial cell wall synthesis (Chapter 20). A number of **ectoenzymes** are present predominantly on the outsides of cell membranes.[329] Enzymes such as phospholipases (Chapter 12), which are present on membrane surfaces, often are relatively inactive when removed from the lipid environment but are active in the presence of phospholipid bilayers.[330,331] The distribution of lipid chain lengths as well as the cholesterol content of the membrane can affect enzymatic activities.[332]

Why are membranes so important to cells? Besides their obvious importance in enclosing and defining the limits of cells, membranes are the result of a natural aggregation of amphipathic molecules. They also represent a natural arrangement for boundaries between different aqueous phases within a cell. In addition, membranes are the "habitat" for many relatively nonpolar molecules formed by metabolism. These include proteins with hydrophobic surfaces and

those with membrane anchors. The semiliquid interior of the membrane permits distortion of the bilayer and the addition or subtraction of proteins (and low-molecular-mass materials) in response to metabolic processes in the adjacent cytoplasm.

The principal factor providing stability to macromolecules and membranes is the hydrophobic nature of reduced organic compounds. This characteristic leads to the separation of lipids, proteins, and other molecules from the aqueous cytoplasm into oligomeric aggregates and membranes. However, the best catalysts, including most enzymes, are soluble in water. Thus, membranes represent thin regions of relative stability adjacent to aqueous regions in which chemical reactions occur readily and which tend to contain the more polar, the smaller, and the more water-soluble materials. The stability of membrane surfaces provides a means of bringing together reactants and of promoting complex sequences of biochemical reactions. For example, membranes contain both oxidative enzymes and reactive, dissolved quinones. The membrane–cytoplasmic interfaces may often be the metabolically most active regions of cells.

Despite their stability, membrane components have a metabolism of their own which is related to the high concentrations of oxidizing enzymes located in or on membranes. Oxidative reactions provide a mechanism for modification of hydrophobic membrane constituents. For example, sterols, prostaglandins, and other regulatory molecules are initially synthesized as hydrophobic chains attached to water-soluble "head groups" (Chapter 21). The hydrophobic products of these synthetic reactions tend to be deposited in membranes. However, attack by oxygen leads to introduction of hydroxyl groups and to a gradual increase in water solubility. As the hydrophilic nature of the compound is increased through successive enzymatic hydroxylation reactions, the hydrophobic membrane constituents eventually redissolve in the water and are completely metabolized. Another process that actively degrades membrane lipids is attack by hydrolytic enzymes such as the phospholipases.

## C. The Transport of Molecules through Membranes

Small neutral molecules, such as water or ethanol, can penetrate membranes by **simple diffusion**.[64,150] The rate is determined by the solubility of a substance in the membrane, by its diffusion coefficient (see Eq. 9-24) in the membrane, and by the difference in its concentration between the outside and the inside of the cell. This concentration difference is commonly referred to as the **concentration gradient** across the membrane. The ease of diffusion through a membrane is described quantitatively by a **permeability coefficient** $P$ which is related to the diffusion coefficient $D$ (Eq. 8-3).

$$J = -D_m K c / \Delta x = -P \Delta c \qquad (8\text{-}3)$$

Here $J$ is the **flux** of molecules across the membrane, i.e., the number of molecules crossing one $cm^2$ per second. $D_m$ is the diffusion coefficient in the bilayer, while $K$ is a **partition coefficient**, the ratio of the concentration of the diffusing solute in the bilayer to that in water. The concentration gradient of the solute across the membrane is $\Delta c$, while $\Delta x$ is the membrane thickness in centimeters. The permeability coefficient $P$ for $H_2O$ through biological membranes[333] is about $1 - 10\ \mu m\ s^{-1}$. For $H^+$ and $OH^-$ $P$ is $0.1\ \mu m\ s^{-1}$. For halide ions diffusing across liposome bilayers $P$ ranges from $10^{-5}$ to $10^{-3}\ \mu m\ s^{-1}$,[334] fast enough to be of some biological significance. However, for most other ions $P$ is less than $10^{-6}\ \mu m\ s^{-1}$. Because of their high lipid content membranes are quite permeable to nonpolar materials. For example, anesthetics usually have a high solubility in lipids, enabling them to penetrate nerve membranes.

### 1. Facilitated Diffusion and Active Transport

While simple diffusion may account for the entrance of water, carbon dioxide, oxygen, and anesthetic molecules into cells, movement of most substances is facilitated by protein **channels** and **transporters**.[335] Genes of 76 families of such proteins have been located in the genomes of 18 prokaryotes.[335a] Some of these provide for **facilitated diffusion**.[169,335] Like simple diffusion, it depends upon a concentration gradient and molecules always flow from a higher to a lower concentration. A distinguishing feature of facilitated diffusion is a **saturation effect**, i.e., a tendency to reach a maximum rate of flow through the membrane as the concentration of the diffusing substance, on the high concentration side, is increased. In this characteristic it is similar to enzymatic action (Chapter 9).

In **active transport** a material is carried across a membrane against a concentration gradient, i.e., from a lower concentration to a higher concentration. This process necessarily has a positive Gibbs energy change (as given by Eq. 6-25) of approximately $5.71 \log c_2 / c_1$ kJ $mol^{-1}$, where $c_2$ and $c_1$ are the higher and lower concentrations, respectively. The transport process must be coupled with a spontaneous exergonic reaction. In **primary active transport** there is a direct coupling to a reaction such as the hydrolysis of ATP to "pump" the solute across the membrane. **Secondary active transport** utilizes the energy of an **electrochemical gradient** established for a second solute; that is, a second solute is pumped against a concentration gradient and the first solute is then allowed to

cross the membrane through an exchange process with the second solute (**antiport** or **exchange diffusion**).

Alternatively, both the first and the second solutes may pass through the membrane bound to the same carrier (**cotransport** or **symport**). Another form of active transport is **group translocation**, a process in which the substance to be transported undergoes covalent modification, e.g., by phosphorylation. The modified product enters the cell and within the cell may be converted back to the unmodified substance. Transport processes, whether facilitated or active, often require the participation of more than one membrane protein. Sometimes the name **permease** is used to describe the protein complexes utilized.

Like facilitated diffusion, active transport depends upon conformational changes in carrier or pore proteins, the equilibrium between the two conformations depending upon the coupled energy-yielding process. Thus, if ATP provides the energy a phosphorylated carrier will probably have a different conformation than the unphosphorylated protein. A carrier with $Na^+$ bound at one site may have a different affinity for glucose than the same carrier lacking $Na^+$. A hypothetical example of the kind of cycle that can function in active transport is provided by the picture of the "sodium pump" given in Fig. 8-25.

## 2. Pores, Channels, and Carriers

To accommodate the rapid diffusion that is often needed to supply food, water, and inorganic ions to cells, membranes contain a variety of small pores and channels. The pores may be nonspecific or they may be selective for anions or cations or for some other chemical characteristics. They may be permanently open or sometimes closed and referred to as **gated**. The gating may be controlled by the membrane electrical potential, by a hormone, by the specific ligand, or by other means. Some pores may be small enough to allow only small molecules such as $H_2O$ to pass through. Others may be large enough to allow for nonspecific simple diffusion of molecules of low molecular mass. Structures are known for only a few.

Large pores tend to be nonspecific, but when the solute approaches the pore diameter in size the specificity increases. Furthermore, diffusion of ions through pores is influenced strongly by any charged groups in or near the pore. Thus, a cation will not enter a pore containing a net positive charge in its surface. Any electrical potential difference across the membrane, resulting from accumulation of excess negative ions within the cell, will also affect the diffusion of ions.[336,337]

***Porins.*** The outer membranes of gram-negative bacteria contain several 34- to 38-kDa proteins known as porins. They form a large number of trimeric pores which allow molecules and ions with molecular masses <600 Da to enter. However, even small proteins are excluded. This appears to be a means for protecting the bacteria against enzymes such as lysozyme (Chapter 12). Four distinct porins are among the most abundant proteins of the outer membrane of *E. coli*.[338] They are designated according to the names of their genes. **OmpF** (a nonspecific, open channel) and OmpC (**osmoporin)** are encoded by *OmpF* (outer membrane protein F) and *OmpC* genes, respectively. **Maltoporin** (LamB) is selective for maltodextrins but also allows other small molecules and ions to pass.[338–341a] Its name comes from its original discovery as a receptor for bacteriophage lambda. PhoE (**phosphoporin)** is a porin with a preference for anions such as sugar phosphates while OmpF prefers cations.[342] Porin **FepA** is ligand gated, opening to take up the chelated iron from ferric enterochelin (Chapter 16). OmpF, OmpC, and PhoE all have 16-stranded β-barrel structures (Fig. 8-20).[199,343] Maltoporin forms a quite similar 18-stranded barrel. Similar porins are present in many bacteria.[199,344]

The porin monomers associate to form trimeric channels as is shown in Fig. 8-20B. They all have a central water-filled, elliptical channel that is constricted in the center to an "eye" ~0.8 x 1.1 nm in size. In this restriction zone the channel is lined with polar residues that provide the substrate discrimination and gating. For example, in OmpF and PhoE there are many positively and negatively charged side chains that form the edge of the eye (Fig. 8-20C). The electrostatic potential difference across the outer membrane is small, but apparently determines whether the porins are in an open or a closed state.[344a] The voltage difference has opposite effects on OmpF and PhoE, apparently as a result of the differing distribution of charged groups.[342,345,346] A key role in determining the voltage dependence may be played by Lys 18 (Fig. 8-20C).[342] Polyamines (Chapter 24), which are present in the outer membrane, induce closing of porin channels by binding to specific aspartate and tyrosine side chains.[347]

The most *abundant* protein in the *E. coli* outer membrane is OmpA. It appears to form a transmembrane helical bundle. Although it is regarded primarily as a structural protein it too acts, in monomeric form, as an inefficient diffusion pore.[350] Mitochondrial outer membranes contain nonspecific pores (mitochondrial porins) that allow passage of sucrose and other saccharides of molecular mass up to 2 to 8 kDa.[351,352] Similar pore-forming proteins have been found in plant peroxisomes.[353]

***Aquaporins.*** Many biological membranes are not sufficiently permeable to water to allow for rapid osmotic flow. For example, the kidney membranes in

portions of Henle's loop have permeabilities as high as 2500 μm $s^{-1}$ (compared to 10–20 μm $s^{-1}$ in a 1:1 cholesterol / phospholipid bilayer).[354-355a] This high permeability is provided by aquaporin-1 (**AQP-1**), formerly called CHIP (channel-forming integral membrane protein) or AQP-CHIP. Aquaporin-1 was first identified in erythrocyte membranes and is present in many tissues. The 28-kDa subunits of the protein form six-helix bundles, each with a pore in the center. These are associated as tetramers in the membrane.[356–358]

Other aquaporins with related sequences occur broadly. There are at least ten in mammals.[359,359a] Plants, which must accommodate to heavy loss of water in hot dry weather, have aquaporins in both plasma membranes and tonoplasts.[360] Bacteria also have aquaporins.[356,361] A defect in aquaporin-2 of the kidney collecting duct leads to **nephrogenic diabetes insipidus**, in which the kidneys fail to concentrate urine in response to secretion of the hormone vasopressin.[355a,362,363]

***Ion channels.*** Most organisms contain a large number of ion channels. One of these, which plays a key role in nerve conduction, is the **voltage-gated $K^+$ channel**. It is closed most of the time but opens when a nerve impulse arrives, dropping the membrane potential from its resting 50 - to 70-mV (negative inside) value to below zero. This voltage change opens the channel, allowing a very rapid outflow of $K^+$ ions.[364–367] The channels then close spontaneously. There are many different $K^+$ channels but most have a similar architecture.[368–370] The three-dimensional structure has been determined for the membrane-spanning part of the $K^+$ channel of *Streptomyces lividans* (Fig. 8-21). The funnel-shaped tetrameric molecule has a narrow **conduction channel** which contains the **selectivity filter**.

The conduction channel is lined largely with hydrophobic groups. The selectivity filter, which discriminates between $K^+$ and $Na^+$, is a short (~1.2-nm-long) narrow (~1.0-nm-diameter) portion of the

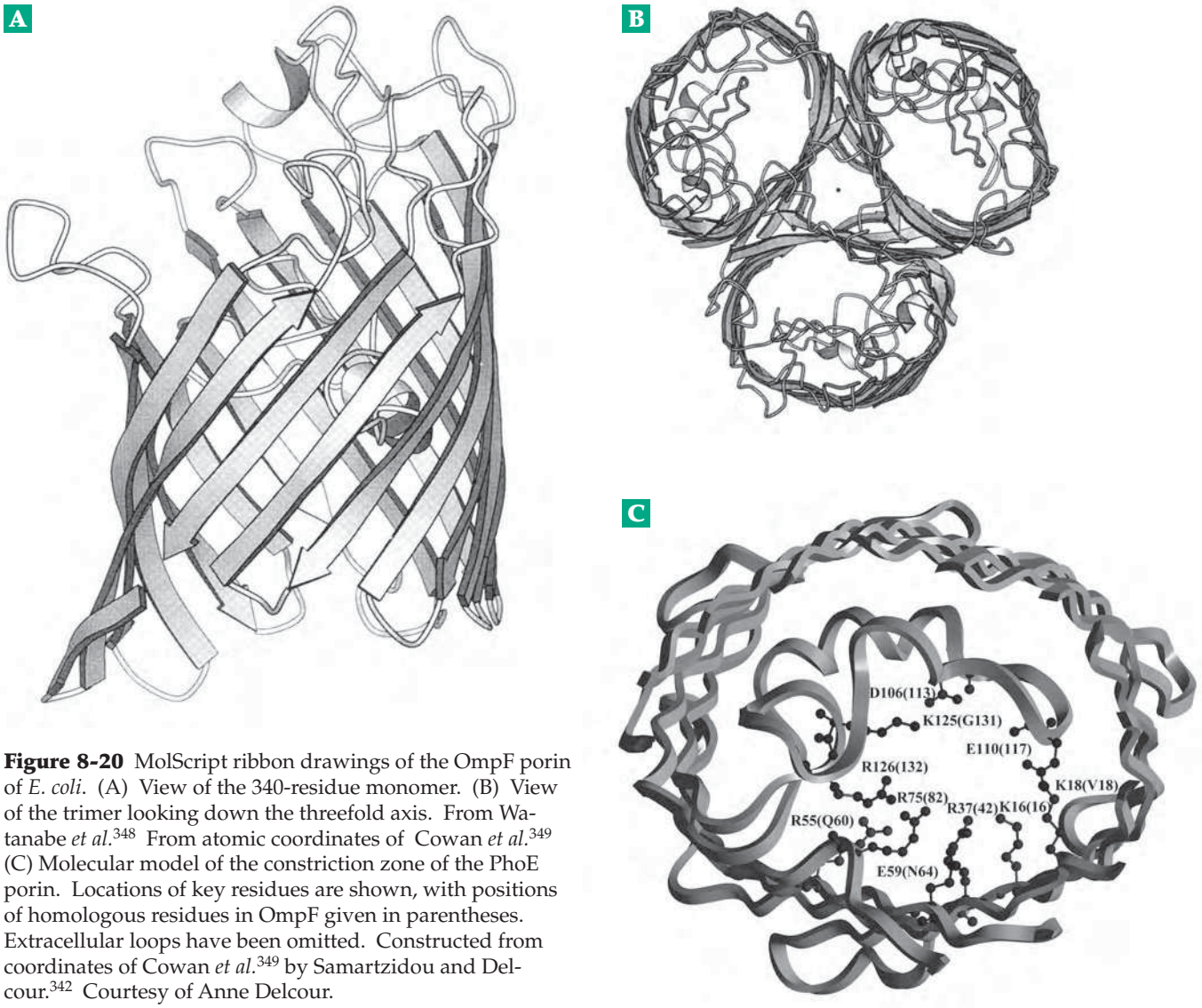



**Figure 8-20** MolScript ribbon drawings of the OmpF porin of *E. coli*. (A) View of the 340-residue monomer. (B) View of the trimer looking down the threefold axis. From Watanabe *et al.*[348] From atomic coordinates of Cowan *et al.*[349] (C) Molecular model of the constriction zone of the PhoE porin. Locations of key residues are shown, with positions of homologous residues in OmpF given in parentheses. Extracellular loops have been omitted. Constructed from coordinates of Cowan *et al.*[349] by Samartzidou and Delcour.[342] Courtesy of Anne Delcour.

pore that is aligned roughly with the center of the bilayer. It is formed by four extended peptide chains, one from each subunit, each having the "**signature sequence**" for $K^+$ channels, TVGYG. In this sequence the peptide carbonyl groups all point into the channel. The consecutive groups of four carbonyls along the channel form binding sites for $K^+$, whose ionic diameter (Table 6-10) is 0.27 nm. The $K^+$ must lose its hydration sphere to fit into the 1.0-nm channel. The site of strongest binding, occupied by $K^+$ in Fig. 8-21C, lies just at the C-terminal ends of four helices, and the partial negative charges of the helix dipoles probably contribute to the binding. The pore may constrict to strengthen the bonds. When a second $K^+$ ion enters the channel it appears to bind ~0.75 nm from the central $K^+$, repelling it and weakening its interaction with the filter, allowing it to pass through the pore. $Rb^+$ (0.30 nm diameter) and Cs (0.34 nm diameter) also pass through.

How are the smaller $Na^+$ (0.19 nm diameter) and $Li^+$ (0.12 nm diameter) excluded from the pore? The pore is too small for the hydrated ions and perhaps too large to bind the dehydrated ion well enough to let it escape from its hydration sphere. Four negatively charged side chains in the cytoplasmic mouth of the pore presumably discourage anions from entering. At the other end, the extracellular entryway is a site that can be blocked specifically by 35- to 40-residue scorpion toxins.[368]

The *S. lividans* $K^+$ channel is *not* voltage gated. Voltage-gating mechanisms must be learned from study of other channels! The voltage-dependent $K^+$ channels from the rat have an $\alpha_4\beta_4$ composition. The α and β subunits coassemble in the endoplasmic reticulum and remain as a permanent complex,[370a–c] After insertion into the plasma membrane the α subunits form a channel as in Fig. 8-21. However, an additional intracellular domain of

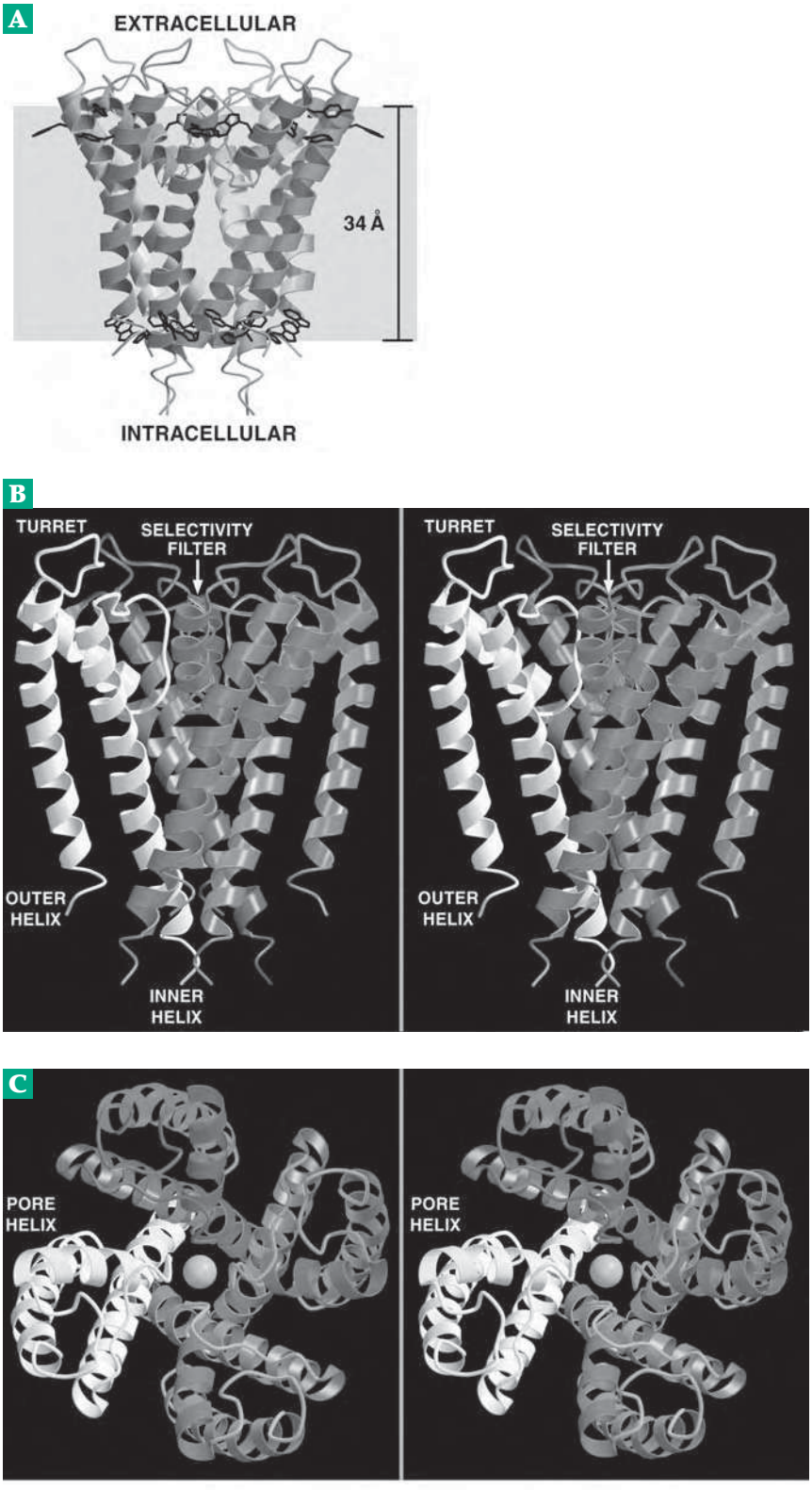



**Figure 8-21** Views of the tetrameric $K^+$ channel from *Streptococcus lividans*. (A) Ribbon representation as an integral membrane protein. Aromatic amino acids on the membrane-facing surface are also shown. (B) Stereoscopic view. (C) Stereoscopic view perpendicular to that in (B) with a $K^+$ ion in the center. From Doyle *et al.*[366]

the α subunits, together with the $\beta_4$ complex, provide an elaborate and as yet poorly understood internal structure. The β subunits are oxidoreductases containing bound NADH, whose function is also uncertain.

There are many more pores and carriers of various kinds in biological membranes. Some, like the *S. lividans* $K^+$ channel, facilitate diffusion of a single ion or compound. They are **uniporters**. Others promote cotransport in which an ion such as $H^+$ or $Na^+$ also passes through the carrier. In some cases a pore is formed by a single molecule. Other pores follow a twofold, threefold, or fourfold axis (Fig. 8-21) of an oligomeric protein.[371] Two different conformations for the protein, one in which a "gate" opens to one side of the membrane and one in which it opens to the other, are usually involved. Interconversion between the two conformations is spontaneous but the equilibrium between them may be influenced by the binding of the solute, by the membrane potential, or by binding of inhibitors or activators. The latter may bind differently to the parts of the carrier (pore) protein exposed on the two sides of the membrane.

***Channel-forming toxins and antibiotics.*** Some of the bacterial toxins known as **colicins** (Box 8-D) kill susceptible bacteria by creating pores that allow $K^+$ to leak out of the cells. One part of the **complement system** of blood (Chapter 31) uses specific proteins to literally punch holes in foreign cell membranes. **Mellitin**, a 26-residue peptide of bee venom,[372,373] as well as other hemolytic toxins and antibiotic peptides of insects, amphibians, and mammals (Chapter 31) form amphipathic helices which associate to form voltage-dependent anion-selective channels in membranes.[374–377]

The polypeptide antibiotic **gramicidin A** consists of 15 nonpolar residues of alternating D- and L-configuration; two molecules can form a channel with a right-handed β helix structure.[378,379] The central 0.48-nm hole in the β helix is large enough to allow unhydrated cations such as $Na^+$ or $K^+$ to pass through. The same peptide, under other conditions, forms left-handed helical channels whose structures are known.[379] Aggregates of $\beta_{10}$ helices may form channels between helices of **suzukacillin**.[380] For this antibiotic as well as for **alamethicin** (Chapter 30) the conductance depends upon the membrane potential, a characteristic shared with the pores of nerve membranes. A variety of synthetic channel-forming peptides have been prepared. Some of these form α-helical bundles with a central pore. A five helix bundle of this type[381] has a structure reminiscent of the acetylcholine receptor channel of neurons.[382]

***Ionophores and other mobile carriers.*** Facilitated diffusion of a molecule or ion is sometimes accomplished by binding to a **mobile carrier**. An example is the diffusion of a complex of $K^+$ with the low-molecular-mass lipid-soluble carrier, or ionophore **valinomycin** (Fig. 8-22). The $K^+$–valinomycin complex diffuses the short distance to the other side of the membrane and discharges the bound ion. If the rates of binding to a carrier and of release from the carrier are greater than those of the diffusion process, Michaelis–Menten kinetics are observed. The maximum velocity $V_{max}$ and Michaelis constant $K_m$ can be defined as in Eq. 9-15 for enzymatic catalysis.

Valinomycin is a **depsipeptide** which contains ester linkages as well as amide linkages. The antibiotic is made up of D- and L-valine, L-lactic acid, and D-hydroxyisovaleric acid. When incorporated into an artificial membrane bathed in a $K^+$-containing medium, valinomycin increases the conductance greatly and when it is added to a suspension of *Streptococcus faecalis* cells the high ratio of $[K^+]_i$ / $[K^+]_o$ falls rapidly.[383] The loss of $K^+$ from cells probably explains the antibiotic activity. However, under suitable conditions, with a high external $[K^+]$, the bacteria will continue to grow and reproduce in the presence of the antibiotic.[384]

Uncomplexed valinomycin has a more extended conformation than it does in the potassium complex.[385,386] The conformational change results in the breaking of a pair of hydrogen bonds and formation of new hydrogen bonds as the molecule folds around the potassium ion. Valinomycin facilitates potassium transport in a passive manner. However, there are cyclic changes between two conformations as the carrier complexes with ions, diffuses across the membrane, and releases ions on the other side. The rate of transport is rapid, with each valinomycin molecule being able to carry ~$10^4$ potassium ions per second across a membrane. Thus, a very small amount of this ionophore is sufficient to alter the permeability and the conductance of a membrane.

Because the stability constant of its complex with potassium is much greater than that with sodium, valinomycin is a relatively specific potassium ionophore. In contrast, the mushroom peptide **antamanide** has a binding cavity of a different geometry and shows a strong preference for *sodium* ions.[388,390] The structure of the $Na^+$–antamanide complex is also shown in Fig. 8-22B. The *Streptomyces* polyether antibiotic **monensin** (Fig. 8-22D),[389,391] a popular additive to animal feeds, is also an ionophore. However, its mode of action, which involves disruption of Golgi functions, is uncertain.[392]

Anions of lipid-soluble phenols such as 2,4-dinitrophenol can serve as effective carriers of *protons* (Chapter 18). However, proteins usually serve as the natural carriers, both of protons and of other ions. A protein is sometimes pictured as rotating to present the solute-binding surface first to one side, then to the other side of a membrane. However, gated pores or channels are probable for most biological transport.

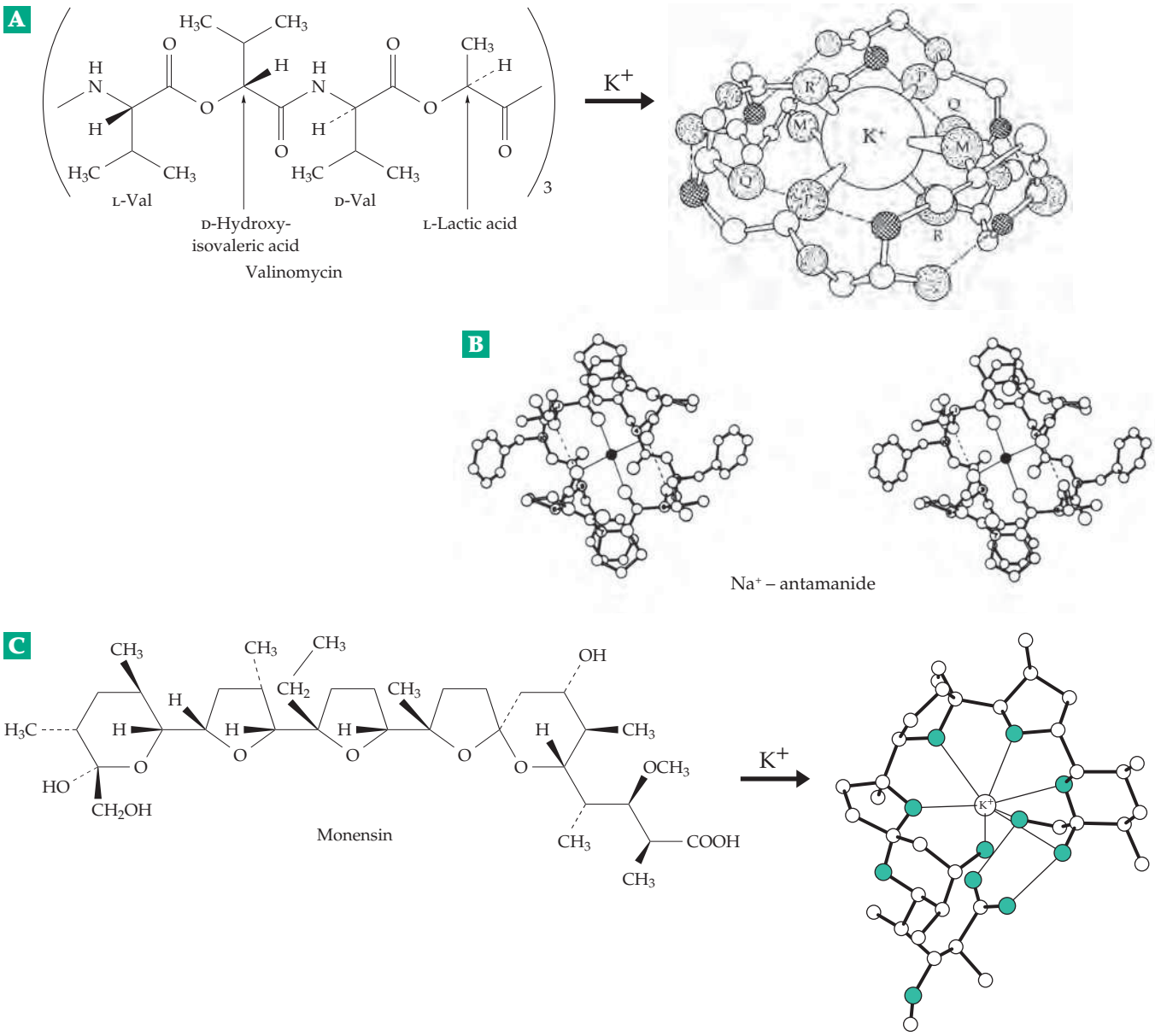



**Figure 8-22** (A) Uncomplexed valinomycin and its complex with $K^+$ (from Duax *et al.*[387]). (B) Stereodiagram of the Na–[Phe[4], Val[6]]antamanide complex. A molecule of $C_2H_5OH$, which forms the fifth ligand to the $Na^+$, is omitted for clarity. From Karle *et al.*[388] (C) Monensin and its complex with $K^+$. From Pangborn *et al.*[389]

## 3. The 12-Helix Major Facilitator Superfamily

A large family of transmembrane facilitators from bacteria and eukaryotes appear to consist largely of 12 transmembrane helices with intervening cytosolic and extracellular loops. Some of these transporters facilitate simple uniport diffusion, but others participate in active transport of the symport or antiport type.[393,394] Several hundred members of the family are known.[395]

***Entrance of sugars into cells.*** It is important that sugars be able to enter cells rapidly. However, the permeability coefficient *P* for D-glucose across a lipid bilayer is only $10^{-6}$ to $10^{-5}$ μm $s^{-1}$. For an intact erythrocyte, *P* is much greater: ~1 μm $s^{-1}$. This is the result of facilitated diffusion by a transport protein with a high specificity for hexose and pentose sugars having a pyranose ring in a C1 chain conformation.[396] This human erythrocyte glucose transporter, now known as **GLUT1**, has a $K_m$ of 1.6 mM for D-glucose but of >3 M for L-glucose. It is a 55-kDa intrinsic membrane glycoprotein migrating in band 4.5 of Fig. 8-14. From the sequence of its cloned gene the unglycosylated carrier was deduced to be a 54-kDa peptide of 492 residues.

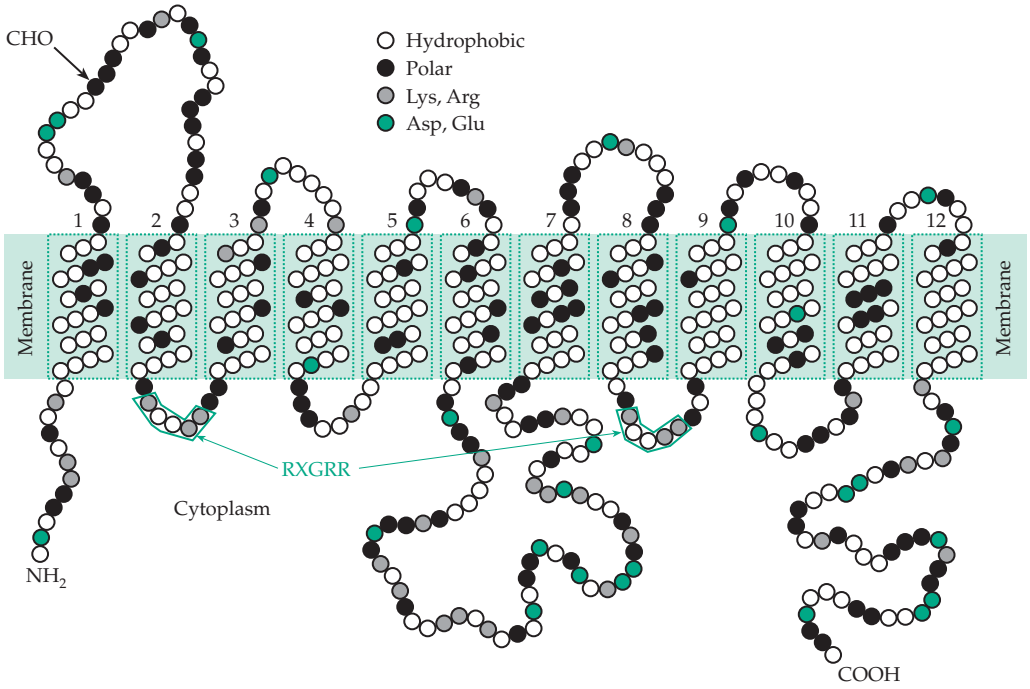
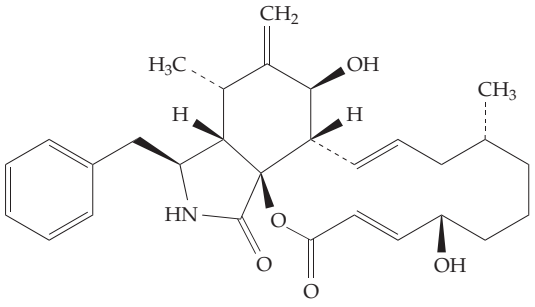



**Figure 8-23** Predicted topology of the human glucose transporter GLUT1. The 12 predicted helices are numbered and the single external N-linked oligosaccharide is marked CHO. The sequence RXGRR (marked) is found in many 12-helix transporters. Its occurrence in these two positions suggests that the transporters may have evolved by duplication of a 6-helix motif. However, the human glucose transporters otherwise show no sequence similarity to other 12-helix transporters. After Bell *et al*.[398] See also Muekler.[400]

The glucose transporter is specifically inhibited by the fungal metabolite **cytochalasin B** which can also be used for photoaffinity labeling of the transporter.[397]




Cytochalasin B

There are at least six closely related facilitative sugar transporters in the human body.[398–400] GLUT1 is found not only in red cells but also in brain[401] and other tissues. GLUT2 is the principal liver transporter and GLUT3 is found along with GLUT1 and GLUT5 in the brain.[401] GLUT5 is primarily a fructose transporter.[402,403] The latter is present in spermatozoa, in which fructose transport is especially important (Box 20-A), and also in the small intestine. Both the GLUT1 and GLUT5 genes are overexpressed in breast cancer.[403] GLUT6 is an unexpressed pseudogene (see Chapter 27) and GLUT7 has been found only in liver microsomes. GLUT4 of skeletal and cardiac muscle and adipose tissue has received a great deal of attention because of its response to insulin[400,404-405a] and to exercise[406] (see Chapter 17).

All of the GLUT family appear to be 12-helix transmembrane-regulated, gated-pore proteins[407,408] with relatively short cytosolic N- and C-terminal ends. The proposed topology of GLUT1, which has been supported by much experimental data,[400,406] is shown in Fig. 8-23. However, the helices are thought to be bundled with the glucose channel centered within a helix bundle. The structure is unknown, but it could resemble the α,α-barrel of glucoamylase (Fig. 2-29), which has 12 helices, a glucose-binding site in the center, and a structure that could easily be modified to form a central pore.

Dehydroascorbate, the oxidized form of vitamin C (Box 18-D) is also transported into cells by GLUT1 and GLUT3.[409] A related transporter carries L-fucose into mammalian cells.[410] Another facilitates the uptake of galactose in yeast.[411]

***Cotransport of sugars and other nutrients with $H^+$ or $Na^+$.*** Epithelial cells of the small intestine or of kidney tubules must take up glucose at low concentra-

tions and discharge it into the bloodstream at a higher concentration.[412] This active transport is accomplished by cotransport of glucose with $Na^+$ in a 1:2 ratio, with the sodium ion concentration gradient across the membrane providing a usable source of energy amounting to 5.8 kJ / equivalent of $Na^+$.[413] The 662-residue human, sodium-dependent transporter **SGLT1** may have 14 transmembrane helices, 5 of which have been proposed to provide the sugar pathway.[414]

Cotransport with $Na^+$ is also observed for transport of many other sugars, amino acids, neurotransmitters, and cofactors.[415] A confusing variety of transporter molecules have been identified and are now being classified into families based on gene sequences as well as function.[416,417] Transporters from intestinal mucosal cells, kidney membranes, and synaptic endings of neurons have been studied most and have been the source for many of the cloned genes.[418] The $Na^+$-dependent transporters of neutral amino acids from these tissues have long been classified as system A, system B, and system ASC (alanine, serine, cysteine) according to substrate specificities.[417,419-420a] Both cationic amino acids and cystine are taken up by kidney tubules and intestinal epithelial cells by another $Na^+$-dependent transporter which is defective in human **cystinuria**, a common metabolic genetic problem.[421,422]

The brain contains several transporters specialized for rapid uptake of neurotransmitters glutamate and aspartate,[423–425] glycine,[426] γ-aminobutyrate (Gaba),[427,428] and catecholamines and also for taurine, L-proline,[429] serotonin,[430] and other substances. Many of these are not only $Na^+$ dependent but also require cotransport of $Cl^-$.[428–430] There are several different glutamate transporter genes with specialized distribution in the brain and other tissues.[423,424]

Cotransport of sugars with $H^+$ is especially common in bacteria[431] but also occurs in eukaryotes. For example, the alga **chlorella** employs hexose / $H^+$ symporters.[432] The most investigated $H^+$ cotransporter is probably the **lactose** (*lac*) **permease** from *E. coli* which enables *E. coli* to take up lactose and other β-galactosides from very dilute solutions.[433–436] From a variety of measurements it has been possible to propose a stacking arrangement for the 12 helices[434] and to identify groups that are essential for function of the 417-residue protein. Glutamates as well as an arginine, and a histidine, all in transmembrane regions, are essential and may be involved in the gating and transport functions.[436] The 469-residue **melibiose permease** of *E. coli* transports α-D-galactopyranosides, including meliobiose (Gal*p*α1 → 6 Glc) and raffinose (Gal*p*α1 → 6 Glc*p*α1 → 2 Fru*f*).[437,438] This permease will couple sugar uptake to either the $H^+$ or $Na^+$ gradient (or to a gradient of $Li^+$ but *not* of $K^+$).[437] Cells of *E. coli* also contain $H^+$ symporters for D-galactose, D-xylose, and L-arabinose that are homologous to the mammalian GLUT proteins.[439]

A series of specific $H^+$-linked cotransporters are found in the brush border membranes of small intestine and in kidney epithelial cells.[440,441] In green plants $H^+$-linked cotransport of amino acids is used in the distribution of amino acids synthesized in the roots and leaves to other parts of the plants.[442,443]

Antiporter or ion exchange transporters are also common. For example, *E. coli* uses a metal ion–tetracycline / $H^+$ transporter to carry the antibiotic tetracycline out of cells. This protein, when present, provides a high level of antibiotic resistance to the bacteria.[444]

## 4. Active Transport Systems

Both bacteria and eukaryotes also possess complex active transport systems for uptake of sugars, amino acids, and other nutrients and for pumping out toxic xenobiotics. One of the most important groups are the high affinity **ABC** (ATP-binding cassette) **transporters**, also called **traffic ATPases**.[445–448a] The *E. coli* genome contains genes for 80 ABC transporters, 54 of which had been identified before the genome sequence was completed.[447] In the human body an ABC transporter enables eukaryotic cells to pump out a large number of different drugs and other foreign compounds. This **multidrug resistance protein** (or P-glycoprotein) not only protects cells but also can seriously interfere with drug treatment. For example, cancer cells with increased amounts of this transporter often arise during chemotherapy. The transporter protein is a single 1280-residue, 170-kDa chain which probably has 12 transmembrane helices and two ATP-binding domains.[448,449] Other human ABC transporters include the antigen processing transporter **TAP** (Chapter 31), the 1480-residue anion transporter **CFTR**, which is defective in cystic fibroses (Box 26-C),[450] the erythrocyte glutathione-conjugate exporter (Box 11-B), and a long-chain fatty acid transporter of peroxisomes. ABC transporters usually consist of four domains. Two are hydrophobic intrinsic membrane domains, each with six membrane-spanning helices and two are peripheral membrane ATP binding domains. All four domains may be in a single peptide chain, as in CFTR, or they may be separate smaller proteins as in bacterial periplasmic permeases.[431,446]

***Periplasmic permeases.*** Gram-negative bacteria contain numerous ABC transporters with components located on the periplasmic surfaces of their plasma membranes. Many of these can be dissociated from the surfaces by **osmotic shock**, i.e., by sudden changes in the osmotic pressure of the medium.[451,452] For example, cells of *E. coli* suspended in 0.5 M sucrose, treated with $10^{-4}$ M EDTA for 10 min, and then diluted

with cold water release ~50 **binding proteins** that hold sugars, amino acids, ions, and other substances tightly with $K_f = 10^6–10^8$ $M^{-1}$.[453,454] An example is the **L-arabinose binding protein** of *E. coli*, a 306-residue peptide organized into two large α / β units with a deep cleft between them. The sugar is bound into this cleft[455,456] by an extensive network of hydrogen-bonding interactions between polar groups in the sugar and in the protein similar to the network that binds either D-galactose (Fig. 4-18) or D-glucose to another of the periplasmic binding proteins.[457,458] Proteins with the same general architecture bind D-ribose, L-arabinose,[458] and maltodextrins.[459] The α1,4-linked maltodextrins, as well as maltose, are major nutrients for *E. coli* and enter the periplasmic space via the previously mentioned LamB porin. Other periplasmic proteins bind histidine,[460] basic amino acids,[461] branched chain amino acids, only leucine, oligopeptides, polyamines,[462] and the tetrahedral anions phosphate[454] and sulfate.[463] The sulfate$^{2-}$ anion is held to its binding protein by seven hydrogen bonds–one from a serine –OH, one from an indole ring NH, and five from peptide NH groups, three of which are at the positive ends of α helices. No permanently charged groups nor cations nor water molecules come into contact with the $SO_4^{2-}$.[463] The phosphate-binding protein also binds a tetrahedral dianion $HPO_4^{2-}$ but it doesn't bind sulfate. It forms numerous hydrogen bonds with its ligand and also an ion pair with a guanidinium group. An aspartate carboxylate hydrogen bonds to the –OH

## BOX 8-D COLICINS: ANTIBIOTIC PROTEINS

Certain strains of *E. coli* and related bacteria synthesize proteins known as colicins that kill cells of other susceptible strains.[a–c] Three kinds of colicins, each encoded in its own small DNA plasmid (colicinogenic factor), are known. **Colicin E3** is a 58-kDa ribonuclease (RNase) that attacks 26S ribosomal RNA of susceptible bacteria;[d] **colicin E2** is a deoxyribonuclease (DNase) that cleaves the bacterial chromosome.[e] **Colicin E1** and its relatives, colicins A, B and Ia, Ib and N, attack the bacterial inner membranes and form lethal pores which allow $K^+$ and other ions to flow out of the cell. The presence of a single channel will kill the bacterium. The effect is similar to that of valinomycin.[f,g] The small colicin V is an 88-residue peptide antibiotic that is secreted by a dedicated ABC export system.[h] Colicins are members of a larger group of **bacteriocins**. One of these proteins, megacin Cx from *Bacillus megaterium*, kills bacteria of sensitive strains by blocking protein synthesis.[i]

The channel-forming colicins bind to bacterial surface molecules that serve as their receptors. For example, colicin N binds to the abundant *E. coli* surface protein OmpF. Interaction with a complex of membrane proteins known as tol Q, R, A, and B then leads to translocation across both outer and inner membranes and refolding of the colicin in a pore-forming conformation.[g] This mechanism is also used by colicins A, E, and K. The single-stranded bacteriophages M13, fd, and f1 "parasitize" the same transport system. Colicins B, D, I, and M enter bacteria with the aid of a second transport system consisting of proteins TonB, ExbB, and ExbD which participate in uptake of chelated iron (Chapter 16) and of vitamin $B_{12}$. This system is also parasitized by bacteriophages T1, T5, and ϕ80.[j]

The N-terminal portion of the 522-residue polypeptide chain of colicin E1 appears to be required for transport into the membrane and the central part for binding to the receptor; the channel-forming property is characteristic of the C-terminal region.[k] A similar organization has been established for the smaller colicin N:translocation domain, (residues 1–66), receptor domain, (residues 67–182), and pore-forming domain (residues 183–387).

The colicin E1 plasmid is a 4.43 MDa circular double stranded DNA molecule consisting of 6646 base pairs.[l] Only one site is susceptible to cleavage by the restriction endonuclease ECoR1 (Chapter 26) This feature has led to its widespread use in cloning of genes.

---


[a] Luria, S. E. (1975) *Sci. Am.* **233**(Dec), 30–37

[b] Parker, M. W., Tucker, A. D., Tsernoglou, D., and Pattus, F. (1990) *Trends Biochem. Sci.* **15**, 126–129

[c] Parker, M. W., and Pattus, F. (1993) *Trends Biochem. Sci.* **18**, 391–395

[d] Escuyer, V., Boquet, P., Perrin, D., Montecucco, C., and Mock, M. (1986) *J. Biol. Chem.* **261**, 10891–10898

[e] Dvhsllrt, K., and Nomura, M. (1976) *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3989–3993

[f] Wiener, M., Freymann, D., Ghosh, P., and Stroud, R. M. (1997) *Nature (London)* **385**, 461–464

[g] Evans, L. J. A., Labeit, S., Cooper, A., Bond, L. H., and Lakey, J. H. (1996) *Biochemistry* **35**, 15143–15148

[h] Fath, M. J., Zhang, L. H., Rush, J., and Kolter, R. (1994) *Biochemistry* **33**, 6911–6917

[i] Brusilow, W. S. A., and Nelson, D. L. (1981) *J. Biol. Chem.* **256**, 159–164

[j] Derouiche, R., Bénédetti, H., Lazzaroni, J.-C., Lazdunski, C., and Lloubès, R. (1995) *J. Biol. Chem.* **270**, 11078–11084

[k] Griko, Y. V., Zakharov, S. D., and Cramer, W. A. (2000) *J. Mol. Biol.* **302**, 941–953

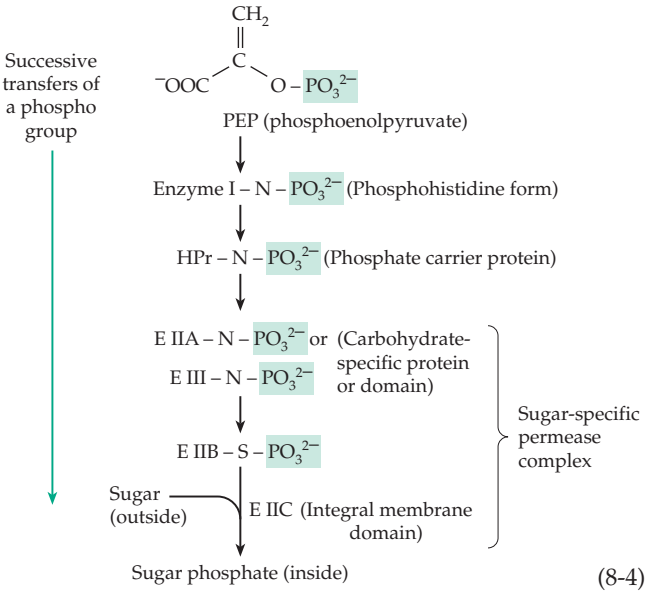
[l] Chan, P. T., Ohmori, H., Tomizawa, J.-I., and Lebowitz, J. (1985) *J. Biol. Chem.* **260**, 8925–8935

of the phosphate. Since the sulfate ion lacks the necessary H to form this bond it is excluded from the binding site.[454]

All of these binding proteins have similar architectures. The ligands fit into a groove between two domains as is seen in Fig. 8-24 for the histidine binding protein. The histidine is held by formation of carboxylate–arginine and $-NH_3^+-$aspartate ion pairs and an additional hydrogen bond to the imidazole. The proteins are able to bend in the hinge region between the two domains to give a better fit to their ligands.

The periplasmic binding proteins function together with the other subunits of the ABC transporter system. One of the best understood systems is encoded by the histidine transport operon of *Salmonella typhimurium*.[445,453] There are four genes: *hisJ* (encoding the histidine-binding protein), *hisQ*, *hisM*, and *hisP*. The Q and M proteins are hydrophic integral membrane proteins which interact with two copies of the P protein, which contains the ATP-binding motif, to form a His $QMP_2$ membrane complex. The soluble HisJ transfers its bound histidine to this complex, which with the hydrolysis of ATP supplying the driving force transfers the histidine across the membrane, presumably via a channel.[445] In *E. coli* the gene *malE* encodes the periplasmic maltose binding protein, while *malF* and *malG* are genes for the integral plasma membrane components. The *malK* gene encodes an ATP-binding protein homologous to the *hisP* protein. As with the histidine permease, a multiprotein complex $MalFGK_2$ is formed. It accepts maltose or a maltodextrin and ATP is hydrolyzed to drive the transport.[459] *MalT* encodes a positive regulatory protein that stimulates transcription of the mal genes.[464] Membrane transport systems are made more complex by the fact that several of the binding proteins serve also as receptors for stimulation of **chemotaxis** (Chapter 19).

***The bacterial phosphotransferase system.*** A third system for uptake of sugars is utilized by *E. coli* and by many other bacteria. This phosphoenolpyruvate-dependent phosphotransferase system converts glucose, mannose, fructose, other sugars, or mannitol into their 6-phosphate esters, at the same time transporting the latter across the membrane (a group translocation).[431,465] Four proteins form a cascade (Eq. 8-4). Phosphoenolpyruvate, an intermediate in sugar metabolism whose high energy of hydrolysis provides the




(8-4)

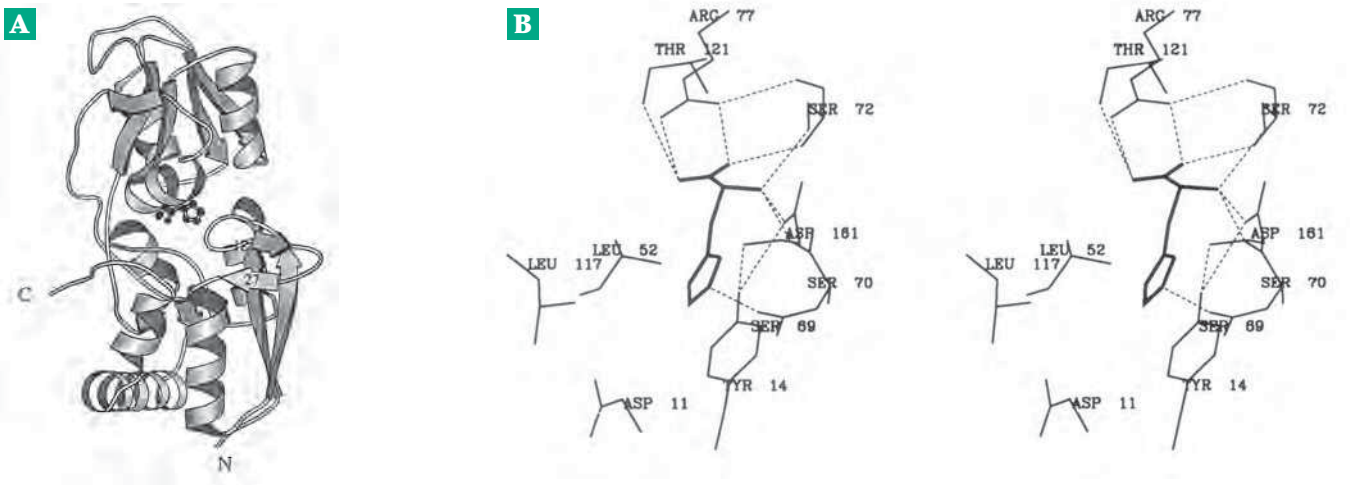



**Figure 8-24** (A) MolScript ribbon drawing of the periplasmic histidine-binding protein HisJ, a component of an ABC transporter system of *Salmonella*. The bound L-histidine is shown as a ball-and-stick model. (B) Stereoscopic view of the histidine-binding site showing hydrogen-bonding interactions of protein side chains with the histidine. From Oh *et al.*[460] Courtesy of Giovanna Ferro-Luzzi Ames.

driving force, phosphorylates $N^{\varepsilon 2}$ of a histidine side chain on **enzyme I**, a large 64-kDa soluble membrane-associated protein.[466-467a]

The phospho group is then transferred sequentially to the small 88-residue dimeric phosphate carrier protein **HPr**, to the carbohydrate-specific membrane proteins IIA and IIB, and to the sugar being transported. A histidine side chain at position 15 of HPr is phosphorylated by PEP to form $N^{\delta 1}$-phosphohistidine.[468] The $NH^{\varepsilon 2}$ of the same histidine forms a hydrogen bond to C-terminal glutamate 85, which is also hydrogen bonded to an arginine side chain. This chain of interacting groups may function in the phospho transfer reactions.[469] Both enzyme I and HPr function in the transport of many ligands but there are specific **enzymes II** (EIIs) for each sugar or other ligand.[470] In *E. coli* there are at least 13 different PTS transporters.[471] A complete EII usually consists of three domains–EIIA (or EIII), EIIB, and EIIC. In some cases all three domains are in a single polypeptide chain, but in other cases they are individual proteins or some combination of individual and bifunctional proteins. The 637-residue mannitol-specific EII contains all three domains. Domains A and B are cytoplasmic while the N-terminal segments form the integral membrane C domain. The two phosphorylation sites are His 554 (EIIA) and Cys 384 (EIIB).[472] The glucose-specific EII from *E. coli* consists of two subunits, IIA and IICB.[471,473–475] The mannose transporter has three subunits representing domains IIAB, IIC, and an additional integral membrane subunit IID.[476] In addition to their direct transport functions, components of the PTS system play regulatory roles in chemotaxis, transcription, and control of other transporters.[477,478]

## 5. Transport of Ions

Cell membranes are impermeable to most ions. Only a small number of ions can enter cells readily and these usually do so with the assistance of protein channels or pores. The principal anion of plasma (Box 5A) is $Cl^-$, which passes through membranes readily by virtue of the presence of channel-forming proteins. Chloride ions are often distributed across membranes passively according to Eq. 8-5, which describes the **Donnan equilibrium**.[167,479,480]

$$[K^+]_i[Cl^-]_i = [K^+]_o[Cl^-]_o \qquad (8\text{-}5)$$

Here the subscripts i and o refer to the inside and the outside of the cell, respectively. The potassium ion concentration within a cell is maintained at a high value by the operation of the $Na^+ + K^+$ pump and by the presence of nondiffusible anions within the cell. According to Eq. 8-5, the internal chloride concentration must be low, with the product of $[K]_i[Cl]_i$ equaling that of the low exterior $[K^+]$ and high exterior $[Cl^-]$. The internal $[Na^+]$ and $[Ca^{2+}]$ are both low, while the internal $[K^+]$ is high. These differences are also linked to the membrane potential (Eq. 8-2), which is ordinarily expressed as a negative voltage of the interior of a cell, mitochondrion, plastid, etc. with respect to a reference electrode in the external medium.

The maintenance of both the membrane potential and the steep gradients of ionic concentrations is essential to cells, both as a means of coupling metabolic energy to transport and other processes and for electrical signaling. The effects are most pronounced for mitochondrial membranes for which $E_m$ may attain –140 to –170 mV and for plasma membranes of excitable cells such as neurons ($E_m$ = –70 to –90 mV). For liver and kidney cells $E_m$ of plasma membranes may be approximately – 35 mV and for erythrocytes only – 9 mV.[480]

In excitable cells electrical impulses are initiated by opening or closing ion channels. They are propagated along an axon by a complex sequence of opening and closing of voltaged-gated channels,[481] a process that is described in Chapter 30. All cells appear to also contain ATP-driven ion pumps as well as simple channels, cotransporter proteins, and ion exchangers.

***Anions.*** Cell membranes have long been known to be relatively permeable to $Cl^-$ and other small anions. However, the molecular basis of this permeability is quite complex.[482] Voltage-gated selective anion channels, often called **chloride channels**, are important in electrically excitable membranes, where they ensure a high resting chloride conductance and stability.[483] The gene for one of these channel proteins was first cloned from the electric ray *Torpedo*[484,485] and is designated *Clc*-0. Similar channels have been found in organisms ranging from bacteria to yeast, green plants, and vertebrate animals. The yeast genome contains just one *Clc* gene[486] but mammals have at least nine. One of these, *Clc*-2, is defective in **myotonia congenita**, a human disease of impaired muscle relaxation; and in similar diseases of mice, goats, and horses.[483,487] A mutation in the chloride channel *Clc*-5 causes kidney problems, including proteinuria, hypercalciuria, and kidney stones.[483]

Erythrocyte membranes permit rapid transport of anions to allow for the exchange of $Cl^-$ within the red cells for $HCO_3^-$ generated by tissue metabolism. The $HCO_3^-$ binds to deoxyhemoglobin (Eq. 7-47) and is carried to the lungs, where it is released upon oxygenation of the hemoglobin. Then the reverse exchange of internal $HCO_3^-$ for $Cl^-$ occurs. This electroneutral ion exchange is mediated by the **band 3 protein** (Fig. 8-14) which contains a channel that allows anions but not cations to pass.[238,239,488] The band 3 protein, also known as **AE1** (anion exchanger 1), is found principally in red blood cells but is also present in kidney tubules.[489]

Related proteins occur in other tissues.[488] The 911-residue band 3 protein consists of two distinct parts of nearly equal size. The N-terminal portion is attached to the membrane skeleton (Fig. 8-16). The C-terminal part, which is embedded in the membrane, is thought to form 14 transmembrane helices and to contain the ion exchange channel or channels.[489a] As previously mentioned, defects in the N-terminal portion cause spherocytosis. The mutation Arg 589 His in the C-terminal half causes **renal tubular acidosis** in which the kidneys do not adequately remove acids from the body.[238,489] Band 3 proteins can also exchange phosphate, sulfate, and phosphoenolpyruvate for $Cl^-$ or bicarbonate.

Another chloride channel, which is regulated by cyclic AMP (Chapter 11), functions in secretory epithelia. Its regulation is faulty in **cystic fibrosis** (Box 26-A), one of the most common human genetic defects, especially among persons of European descent.[490] As was previously mentioned, this **cystic fibrosis transmembrane conductance regulator** (CFTR) is a member of the ABC superfamily of transporters. The large 1480-residue protein apparently has two 6-helix membrane-spanning domains, two cytoplasmic nucleotide-binding domains,[491] and another large cytoplasmic regulatory domain.

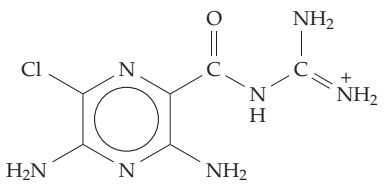
In addition to AE1 (band 3 protein), kidneys depend upon other modes of reabsorption of $HCO_3^-$ from the proximal tubules. These include a $Na^+/HCO_3^-$ cotransporter, which seems to be related to the AE family of ion exchangers. However, it transfers three $HCO_3^-$ ions per $Na^+$ and is therefore highly **electrogenic**.[492,493] Transporters for phosphate, sulfate, and small organic anions are found in many organisms. Bacterial periplasmic transporters have already been described. Plants employ $H^+$/phosphate,[494] $H^+$/sulfate,[495] and $H^+$/nitrate[496] cotransporters. Phosphate transporters are probably essential to all organisms. One of the best known is the mitochondrial $P_i/H^+$ cotransporter which carries phosphate ions originating from hydrolysis of ATP to ADP + $P_i$ back into the mitochondria.[497,498] See also Table 18-8. A human $Na^+/P_i$ cotransporter in the kidney is also essential. An X-linked trait leading to inadequate synthesis of the transporter causes hypophosphatemic vitamin D-resistant **rickets**.[499]

Monocarboxylates such as lactate and pyruvate enter animal cells with the aid of monocarboxylate/$H^+$ cotransporters of low specificity.[500] A $Cl^-$/oxalate transporter is one of several ion exchange proteins in the kidney.[501] Transport systems for ADP, phosphate, dicarboxylates, and other anions are very active in mitochondrial membranes (Chapter 18).

***Cation channels.*** When a nerve impulse passes along an axon gated pores or channels specifically permeable to $Na^+$ and $K^+$ open for short periods of time as a result of changes in membrane potential induced by the advancing wave of the action potential (Chapter 30). We have already considered the structure of a potassium ion channel (Fig. 8-21). However, there are at least 30 types of $K^+$ channels that can be distinguished.[502] Many of these appear to have similar channel structures but to serve a variety of purposes. While the $K^+$ channels of neurons are voltage gated many others are controlled by hormones, neurotransmitters, or mechanical stimuli.[503] One of the most investigated $K^+$ channels, known as $K_{ATP}$, sets the resting potential in the insulin-secreting β cells of the pancreas by facilitating a flow of $K^+$ *into* cells. Such channels, which help equilibrate intracellular and extracellular $K^+$ at near equilibrium, are found in cells with low negative values of $E_m$. They are called *inwardly rectifying*. When the internal glucose concentration in the β cells rises it initiates a complex signaling sequence involving blockage of the $K_{ATP}$ channels by ATP, opening of voltage-sensitive $Ca^{2+}$ channels, and insulin secretion.[504–506] The $K_{ATP}$ channel is also blocked by **sulfonylureas**. As a result, these compounds induce insulin release and are useful in treatment of diabetes meltitus (Box 17-G). When its gene was cloned the **sulfonylurea receptor** was found to be a transmembrane protein of the ABC transporter family.[505] The potassium channel protein is another subunit of the transporter. Similar $K_{ATP}$ channels are found in the kidneys[507] and also in embryonic cells in which they may participate in regulation of the cell cycle.[508]

Voltage-regulated **sodium channels** are the major participants in propagation of nerve impulses. The large 260-kDa α subunit of the sodium channel of nerve membranes contains four homologous repeat sequences, each of which may form transmembrane helices and also contain a loop that may participate in forming a pore similar to the $K^+$ pore of Fig. 8-21.[509-510a] However, the structure is uncertain.[511] The channel complex also contains 36- or 33-kDa $\beta_1$ and $\beta_2$ subunits that appear to be members of the Ig superfamily.

Epithelial cells contain a quite different $Na^+$ channel that participates in reabsorption of urinary $Na^+$ and in control of blood pressure (Box 22-D). The channel consists of at least three structurally similar subunits, each with two predicted transmembrane helices and a large extracellular domain.[512] These channels are blocked specifically by the diuretic compound **amiloride**.[513]




Amiloride, a diuretic

The best known sodium channel is, in fact, a general cation channel which is part of the "nicotinic" **acetylcholine receptor**. Acetylcholine is the principal excitatory transmitter in the peripheral nervous system and upon release from synaptic endings or neuromuscular junctions occupies receptors on the "postsynaptic" membranes of one or more adjacent neurons. When acetylcholine binds in the ion pore, the acetylcholine receptor opens and cations flow out, depolarizing the membrane. Under favorable circumstances, this initiates a nerve impulse (action potential) in the postsynaptic neuron. The receptor gene was first cloned from *Torpedo* and the receptor protein has been studied extensively.[382] The 290-kDa protein consists of five similar-sized subunits with an $\alpha_2\beta\gamma$ structure. These form a nearly symmetric fivefold oligomer with a pore in the center. Images of both the open and the closed states, obtained by electron microscopy at a resolution of 0.9 nm, suggest that the inner pore of ~2.6 nm diameter is formed by five α helices, one from each subunit. It is open when acetylcholine binds to the two α subunits and closes to a much smaller diameter when the acetylcholine leaves (and is destroyed by hydrolysis).

**Calcium channels** are a third major group of cation-selective channels.[514] As pointed out in Box 6-D, calcium ions are involved in a very wide range of signaling functions. These are discussed in several places in this book. Several of these functions depend upon voltage-gated $Ca^{2+}$ channels. Muscle is rich in **L-type** or DHP-sensitive channels (Box 6-D) which play a role in transmission of nerve impulses to muscles by allowing rapid flow of calcium ions into cells from outside.[515]

The structure appears to be homologous to that of voltage-gated $Na^+$ channels with a large 170-kDa subunit with a fourfold repeat plus smaller subunits. The **ryanodine receptors** of muscle control the release of $Ca^{2+}$ from stores in the endoplasmic reticulum[516,517] (see also Chapter 19). These receptors are ligand gated, being activated by cyclic ADP-ribose (Chapter 11). Two additional types of voltage-sensitive $Ca^{2+}$ channels, **N** and **P**, are found in the central nervous system.[514,518] Another ligand gated calcium channel has been found in endothelial cells.[519] It is activated by **sphingosylphosphocholine**[519] rather than by cAMP-ribose or inositol triphosphate.[520] A sodium / calcium ion exchanger and cotransporter[521] utilizes the $Na^+$ gradient to exchange three external $Na^+$ ions for one internal $Ca^{2+}$.

Many aspects of calcium function are poorly understood. Among these is the role of a group of proteins known as **annexins** (formerly lipocortins, calpactins, endonexins, etc.).[522] The ten or more members of the annexin family[523–527] share the property of binding to phospholipid membranes in the presence of $Ca^{2+}$. One of the several proposed functions of annexins is formation of $Ca^{2+}$ channels, a function that is suggested by the modular three-dimensional structures.[525,526] Other suggested functions include roles in membrane fusion, exocytosis, and adhesion.

***Active transport of cations.*** Most organisms take up ions from their surroundings by active transport. Green plants extract essential nutrients from the extremely dilute solutions in contact with their roots. Microorganisms such as yeast and bacteria have the same ability, and specific concentrating systems for many ions such as $K^+$, $Ca^{2+}$, sulfate, and phosphate have been identified. The skin of a frog can take up $Na^+$ from a $10^{-5}$ M solution of NaCl and extrude it into the internal fluids whose $Na^+$ concentration may be greater than 0.1 M. Ions can also be concentrated from internal fluids and excreted at higher concentrations. Some seabirds and marine animals rid their bodies of excess salt by secretion from salt glands. The lining of the human stomach is able to concentrate hydrogen ions in gastric juice to ~0.16 M.

Organelles within cells have their own ion-concentrating mechanisms. Thus, mitochondria can concentrate $K^+$, $Ca^{2+}$, $Mg^{2+}$, and other divalent metal ions as well as dicarboxylic acids (Chapter 18). The entrance and exit of many substances from mitochondria appear to occur by exchange diffusion, i.e., by secondary active transport. Such ion exchange processes may also occur in other membranes.

***ATP-driven ion pumps.*** Within virtually all cells the sodium concentration is relatively low, while that of potassium is high (Box 5-A). One theory[528] regards the cytoplasm as analogous to an ion exchange resin with fixed charges in a lattice. Highly crosslinked ion exchange resins exhibit specificity toward binding of certain ions; e.g., sulfonic acid resins tend to bind $K^+$ preferentially, while phosphonic acid resins tend to bind $Na^+$. Do proteins also prefer $K^+$ to $Na^+$?

In contrast to the ion exchange theory, much evidence indicates that cells have an active **ion pump** that removes $Na^+$ from cells and introduces $K^+$. For example, the cytoplasm of the giant axons of nerves of squid can be squeezed out and replaced by ionic solutions. Erythrocyte ghosts can be allowed to reseal with various materials inside. Ion transport into or out of cells has been demonstrated with such preparations and also with intact cells of many types. Such transport is blocked by such inhibitors as cyanide ion, which prevents nearly all oxidative metabolism. However, the cyanide block can be relieved by introduction into the cells of ATP and other phosphate compounds of high group-transfer potential.

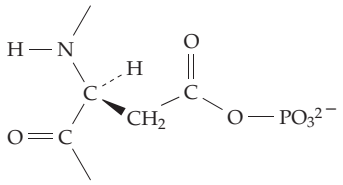
Uptake of $K^+$ by cells and extrusion of $Na^+$ from cells are also specifically blocked by "cardiac glycosides" such as **ouabain** (Fig. 22-12). Ouabain labeled with $^3H$ binds to the outer surface of cells, and from

this binding it was estimated that erythrocytes possess 100–200 ion pumping sites per cell (~1 site / $\mu m^2$).[529] For the HeLa cell (a widely studied strain of human cancer cells) $10^5$ to $10^6$ sites / cell (~10 / $\mu m^2$) were found. Further experiments showed that in the presence of $Na^+$ within the cell and $K^+$ on the outside of the cell, ATP is hydrolyzed. The rate of hydrolysis was directly related to the concentrations of these two ions and to the number of ouabain binding sites and also required the presence of $Mg^{2+}$. These observations led to the concept of an **($Na^+$ + $K^+$)- activated ATPase** (often abbreviated **$Na^+$, $K^+$-ATPase**) as synonymous with the membrane-bound ion pump. Within the cell $Na^+$ must be located on one side of the membrane and $K^+$ on the other to activate this enzyme. However, the purified enzyme would be expected to hydrolyze ATP in the test tube in the presence of $Na^+ + K^+ + Mg^{2+}$. Such a protein was isolated from several sources and has been studied intensively. It is an αβ mixed dimer with molecular masses of ~113 kDa for the α chains and ~55 kDa for the glycoprotein β chains.[530,531] The proteins may associate to $\alpha_2\beta_2$ tetramers in membranes. The genes for various isoforms of the proteins from several sources have been cloned and sequenced.[531,532] The large α subunit may span the bilayer of the membrane as many as ten times; the glycoprotein β subunit is thought to be largely on the outer surface and may have only one membrane-spanning helix.[533,534] A small 68-residue γ subunit copurifies with the pump protein. It may be involved in control of the ATPase, which has complex regulatory properties.[535,536] Sulfatides (ceramide galactose-3-sulfate) may also play a role in the enzymatic activity.[535]

The sodium–potassium pump displays a curious stoichiometry. *Three sodium ions are pumped from the inside and two potassium ions from the outside of a cell for each molecule of ATP cleaved*. Thus, an excess of positive ions is pumped out with the result that a negative charge develops inside the cell and a positive charge accumulates on the outside. This action of the $Na^+$, $K^+$-ATPase is the primary source of the membrane potential for most eukaryotic cells and is said to be **electrogenic**. Because the cell membrane is somewhat permeable to $K^+$, outward diffusion of $K^+$ through the "leaky" membrane along its concentration gradient helps to maintain the membrane potential as does inward leakage of $Cl^-$. At the same time, $Na^+$ diffuses inward, aided by the membrane potential. Even though the permeability of $Na^+$ is low, a steady state is reached at which the rate of passive inward diffusion of cations just balances the membrane potential set up by the active transport.

The energy for transport of $Na^+$ and $K^+$ by the ion pump is supplied by ATP. The $Na^+$,$K^+$- ATPase does not merely catalyze the hydrolysis of ATP but also couples its cleavage to the pumping of the ions. The pumping of sodium and potassium ions is one of the most important energy-requiring activities of cells. It is said to account for 23% of the ATP utilization in a resting human. *Thus, it constitutes an important fraction of the basal metabolic activity.*

The $Na^+$,$K^+$- ATPase is one of a family of over 50 ion pumps that are characterized by transfer of a phospho group from ATP to an aspartate side chain carboxylate in the invariant sequence DKTG to give an intermediate phosphoenzyme + ADP.[530,537]




Phosphoenzyme intermediate of $Na^+$,$K^+$- ATPase

These **P-type ATPases** are characterized by phosphoenzyme intermediates, by a conserved consensus sequence, and through inhibition by vanadate ion.[537–539] The structures are poorly known. Some consist of single chains (perhaps dimerized) and some have more than one chain. However, the major subunit always appears to have about ten transmembrane helices with a large ~430-residue cytoplasmic domain between the fourth and fifth helices. This domain contains the ATP binding site and the phosphoaspartyl group of the phosphoenzyme.[534] This is Asp 369 for the $Na^+$,$K^+$- ATPase.

In addition to the $Na^+$,$K^+$- ATPases there is a very active **$Ca^{2+}$-ATPase** which transports two $Ca^{2+}$ from the inside of cells to the outside while returning two $H^+$ from outside per ATP.[540–543a] This is the primary transporter by which cells maintain a low internal [$Ca^{2+}$]. During its action it becomes phosphorylated on Asp 351. However, in neurons, in which the membrane potential is maintained at a high negative value by the sodium pump, an $Na^+$/ $Ca^{2+}$ ion exchange plays an even more important role.[540]

Other P-type ATPases include the gastric $H^+$, $K^+$- ATPase, which acidifies the stomach and has a high degree of sequence homology with the $Na^+$, $K^+$- ATPase.[544] Secretion of HCl into the stomach apparently involves diffusion of $K^+$ together with $Cl^-$ from the bloodstream through the cells lining the stomach. The $K^+$ is then pumped back into these cells in exchange for $H^+$ by $H^+$,$K^+$-ATPase.[545] The chloride channel may be in the same protein as the ($K^+ + H^+$) pump.[546] The kidney is the principle acid excretory organ of the body and as such also contains proton pumps. An electrogenic $H^+$- ATPase pumps $H^+$ alone outward through the plasma membranes of fungi and of green plants.[547,548] The resulting proton gradient may be used to provide energy for transport of other materials into cells. A group of metal ion P-type transporters carry copper and other nutrient ions into cells and

extrude $Cd^{2+}$ and other toxic ions.[539] Some **alkaliphilic bacteria** pump $Na^+$ to create a sodium ion gradient.[539a] All of these ion pumping systems require MgATP as the source of energy and function via phosphoenzyme intermediates. The ionic gradients generated can be used to move other ions or nonionic compounds into or out of cells by exchange or cotransport processes. For example, internal $H^+$ may be exchanged for external $Na^+$ in an exchanger-mediated process that assists in control of cytoplasmic pH.[549,549a] The reverse process in *E. coli*[549b] and many other bacteria provides the principal mechanism by which those cells export $Na^+$. This exchange is driven by the electrochemical gradient of the $H^+$ ion created by oxidative phosphorylation (Chapter 18). The $Na^+$ ion, in turn, can be used by bacterial cells to drive other uptake processes, e.g., sugar or amino acid-$Na^+$ cotransport.

What is the mechanism by which ATPase transporters function? We still do not know.[550] The pumping cycles for the $Na^+$,$K^+$- ATPase and the $Ca^{2+}$- ATPase are similar although different in details. The ATPases are reversible and with suitable ionic gradients will work as ATP synthases.[551] A strictly hypothetical model for the $Na^+$, $K^+$-ATPase is shown in Fig. 8-25. There are at least two conformations of the ion pump proteins.[552,552a] In one conformation the protein binds three sodium ions tightly, while in the other conformation it binds two potassium ions. The ATP operates the "motor" that carries out the conformational changes. In Fig. 8-25 the ion pump, in conformation A, is shown embedded in a membrane. In the center, perhaps between three or more transmembrane helices, there is a narrow cavity, perhaps resembling that of the $K^+$ channel (Fig. 8-21), into which chelating groups (e.g., C=O groups of the peptide chain) protrude. These groups form the three binding sites for the 0.19-nm-diameter $Na^+$ ion. The spontaneous binding of the sodium ions triggers a phosphorylation reaction by which a phospho group from the $MgATP^{2-}$ complex is transferred to the side chain carboxyl of the active site aspartate. This phosphorylation in turn triggers a change to the second conformation in which the channel to the outside is open and that to the inside is closed. At the same time the affinity for $Na^+$ is decreased over 100-fold and the sodium ions dissociate on the outside. The affinity for $Na^+$ may decrease because the diameter of the pore is increased to accommodate the larger 0.27-nm diameter of $K^+$ ions, perhaps by a twisting motion of the peptide chains that form the channel.

The next step is loading with two $K^+$ ions. The affinity for $K^+$ in the second conformation is high. The return to conformation 1 with release of $K^+$ to the inside is triggered by hydrolytic removal of the phospho group as inorganic phosphate ($P_i$). It may seem surprising that a channel could be opened and closed so readily with synchronous changes in the number and specificity of ion binding sites. However, recall the type of structural alteration occurring upon oxygenation of hemoglobin (Fig. 7-25). Rotation of the hemoglobin subunits with respect to one another causes small changes in the geometrical relationships of groups protruding into the central channel. This strongly

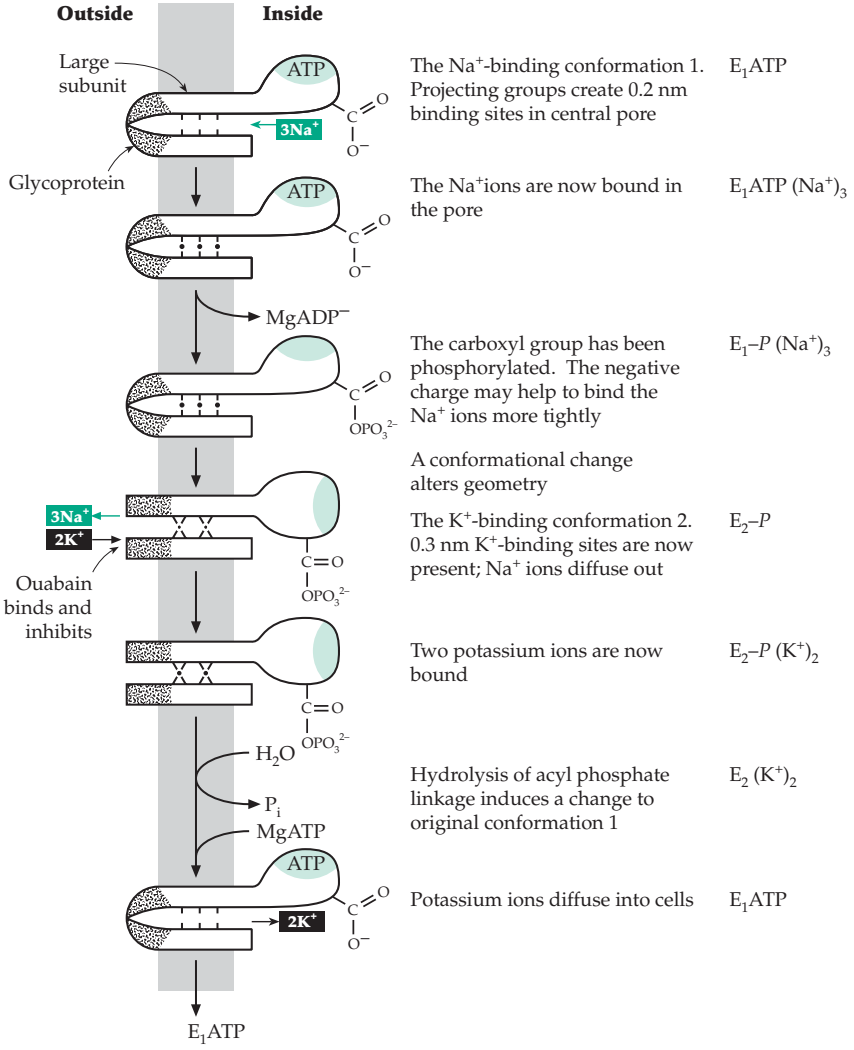



**Figure 8-25** A strictly hypothetical model of a $Na^+$ + $K^+$ pump which operates by ATP-driven opening and closing of a channel at opposite ends and with alternate tight binding of $Na^+$ and $K^+$.

affects the binding of 2,3-bisphosphoglycerate. Very small movements could open up the $Na^+$ binding groups and create new binding sites for the larger $K^+$ ion, using in part the same chelating groups.

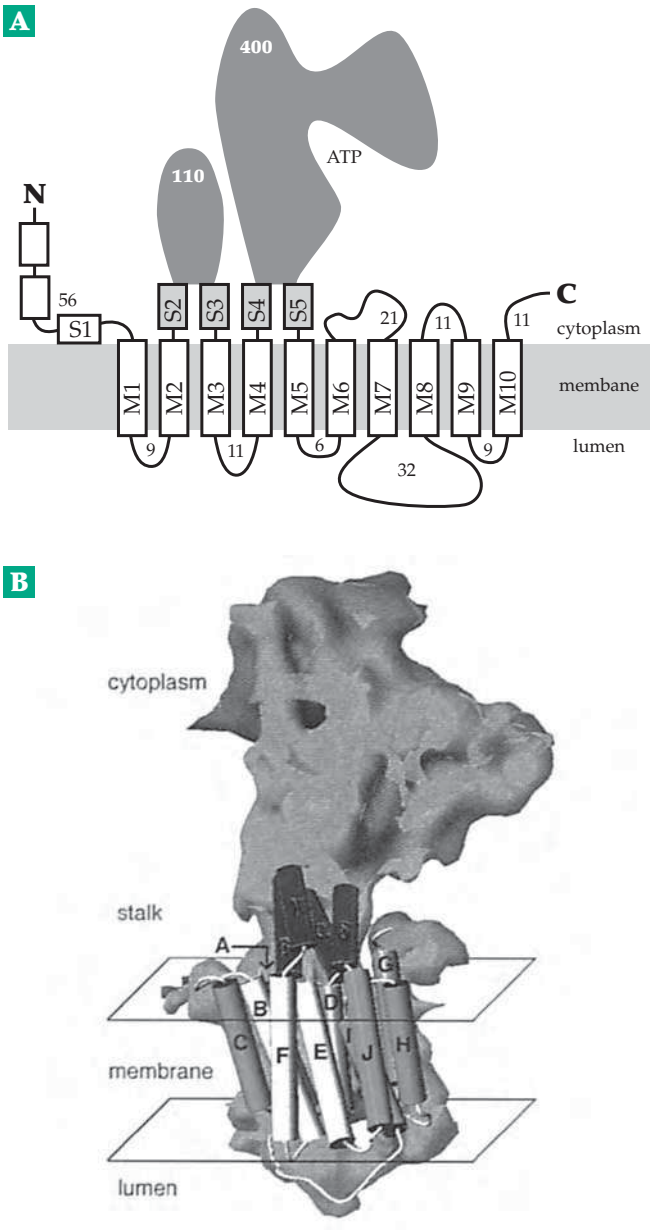



**Figure 8-26** The structure of the 994-residue $Ca^{2+}$-ATPase of the endoplasmic reticulum of rabbit muscle at 0.8-nm resolution. (A) Predicted topology diagram organized to correspond to the electron density map prepared by electron crystallography of frozen-hydrated tubular crystals. The number of amino acid residues in each connecting loop is marked. (B) The electron density map with the predicted structure embedded. The relationships of the helices in (B) to those in (A) are not unambiguous. The helices marked B, D, E, and F in (B) may form the $Ca^{2+}$ channel. The large cytoplasmic loops, which are black in (A), were not fitted. From Zhang *et al.*[553] Courtesy of David L. Stokes.

Images of both the $Ca^{2+}$-ATPase (Fig. 8-26)[553] and the $H^+$-ATPase of *Neurospora* plasma membranes[548] at 0.8 nm resolution reveal similar transmembrane regions and large cytoplasmic domains which are somewhat differently organized. The picture in Fig. 8-26 has been greatly clarified by determination of the structure by X-ray diffraction to a resolution of 0.26 nm.[553a,b,c] Two calcium-binding sites have been located in the transmembrane domain between the helices marked M4, M5, M6, and M8. The $Ca^{2+}$ ions are apparently coordinated by side chains of Asp, Glu, Gln, and Thr. There are three cytosolic domains. The site of phosphorylation, Asp 351, lies within a large ~ 27 kDa **P** (phosphorylation) domain adjacent to the membrane. The ATP is held by a nucleotide-binding **N** domain which must at some point in the cycle move close to Asp 351 for phosphorylation to occur. The third cytosolic (**A**, actuator) domain is thought to be involved in control of the conformational alterations. The nucleotide-binding domain lacks the "P loop" characteristic of many ATPases and GTPases (see p. 648), but is homologous in its sequence to L-2-haloacid dehalogenase (Eq. 12-2)

Two other types of proton-pumping ATPases are considered in Chapter 18. One is the mitochondrial $F_1F_0$ ATPase, which ordinarily operates in the reverse direction as the body's principal **ATP synthase**. The other type, which in some ways resembles the mitochondrial $F_1F_0$ ATPase, is the **vacuolar ATPase** (V-ATPase). These are true proton pumps which acidify vacuoles of plants and also lysosomes and phagocytic vacuoles.[554,555] They are also considered in Chapter 18.

## 6. Exocytosis, Endocytosis, and the Flow of Membrane Constituents

Observation of cells under the microscope with time-lapse photography reveals that the plasma membrane as well as the mitochondria and other organelles are in a constant state of motion. Mitochondria twist and turn and the surface membrane undulates continuously. Vesicles empty their contents to the outside of the cells, while materials are taken into cells through endocytosis. In addition, chemical evidence indicates a directed flow of the materials of which membranes are constructed from the endoplasmic reticulum (ER) to the Golgi vesicles, excretion granules, and plasma membrane (see Fig. 10-8). Along this route new materials are inserted from the cytoplasmic side of the membrane, while enzymes within the vesicles add glycosyl units and make other modifications. The plasma membrane surface area grows quite rapidly. In secretory cells fusion of secretion granules with the plasma membrane also adds additional material to the membrane.

Counterbalancing this expansion of the plasma

membrane is active endocytosis of fluids and of solid materials from outside the cell. This not only brings new materials into the cell but also accomplishes removal of material from the plasma membrane and partial recycling of its components. One form of endocytosis is seen with the ameba.[555a] The cytoplasm flows around a smaller organism or other particle of food enclosing it in an internal membrane-bound compartment (**endocytic vacuole** or **endosome**). This vacuole then fuses with lysosomes which supply the necessary enzymes to digest the food. In a similar way phagocytic cells of our bodies engulf microorganisms or other particles to form **phagosomes** which, over a period of more than 24 hours, undergo extensive biochemical changes.[556] They acquire digestive enzymes, vacuolar ATPase,[557] other proteins needed to kill bacteria, and other parasites.

Uptake of smaller particles including protein molecules occurs by **micropinocytosis**, a process that can be seen only by electron microscopy. This often takes place via **coated pits**, indentations of ~0.3 µm diameter underlain by a thickened membrane.[558–559b] The pit membrane is also coated with protein molecules and appears to have many short bristles or spikes protruding into the cytoplasm (Fig. 8-27A). After endocytosis the coated pits become **coated vesicles** of 0.15–0.25 nm (Fig. 8-27B). Within a few seconds, however, these vesicles lose their coat and become endosomes.

The major protein making up the coat is the 180-kDa **clathrin**, but smaller 33- to 36-kDa peptides of several types also contribute.[561] The coat forms a "basket" with pentagonal and hexagonal faces surrounding the lipid bilayer of the vesicle. At each vertex of the basket is a "triskelion," a trimer of clathrin together with an equal number of the smaller chains. The smallest baskets consist of 12 pentagons plus 4, 8 or more hexagons, a relationship that allows formation of a variety of larger baskets.[562,563] Additional 50- and 100-kDa accessory proteins form a shell around the clathrin cage. That clathrin is essential for normal cell growth has been established by the observation that deletion of its structural gene from yeast is lethal.[564] The addition of more trimer units from a reserve of soluble clathrin in the cytoplasm allows the vesicles to develop and break off from the membrane. From studies with inhibitors it is evident that metabolic energy is required to drive the process.

Other vesicles are surrounded by nonclathrin membrane coats. Some of these originate from **caveolae** (little caves), which act in endocytosis, exocytosis, and transmembrane signaling.[564a,b,c] A **coatomer** complex of eight subunits with molecular masses from 20 to 60 kDa coats vesicles involved in transport between compartments of the Golgi.[565–567]

What is inside a coated vesicle? Cells take up a variety of peptide hormones and proteins. This usually occurs with the aid of specific receptor proteins located in or on the outside of the plasma membrane. Some of these, e.g., receptors for the low-density lipoprotein of plasma (Chapter 22), are clustered in coated pits. Other receptors, such as those for insulin or epidermal growth factor, are spread more evenly across the membrane but collect in coated pits when the hormone binds. Endocytosis provides a means for the cell to take up and in some cases destroy the hormone or the receptor or both.

Transmembrane proteins, including hormone receptors, are incorporated into coated vesicles with the help of **clathrin adapter proteins** (APs). These

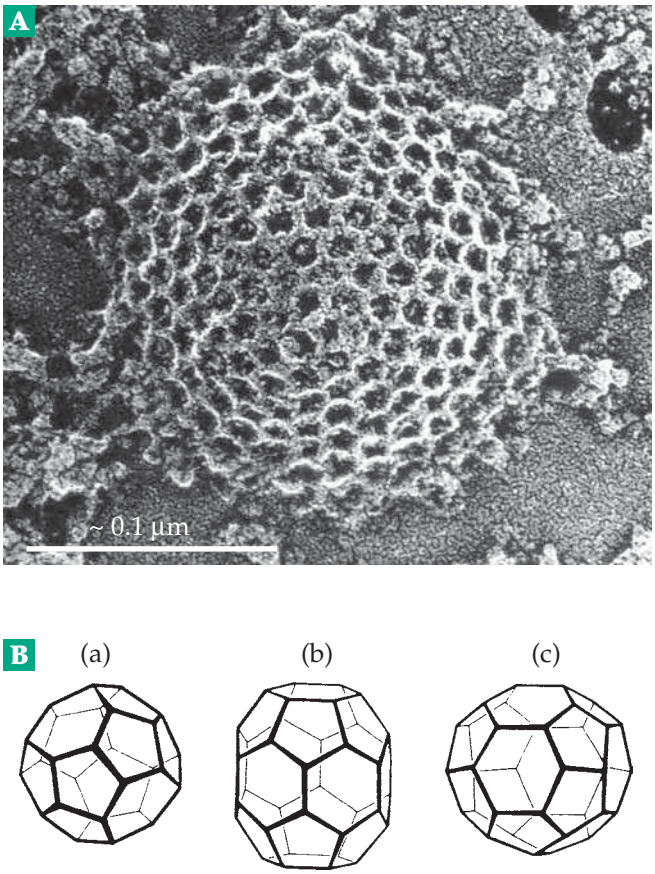



**Figure 8-27** (A) Region of a coated membrane from fibroblasts at an intermediate stage of the budding process, demonstrated by deep etching and rotary replication (by J. E. Heuser[559b]). From Pearse and Bretscher.[560] Courtesy of Barbara Pearse. (B) Three structures identified among the smallest coated vesicles. Structure (*a*) contains 12 pentagons and four hexagons, the latter lying at the vertices of a tetrahedron; (*b*) has a barrel shape built of 12 pentagons and eight hexagons; structure (*c*) also has twelve pentagons and eight hexagons, but the latter are arranged in two arcs of four, related in the same way as the two parts of a tennis ball. Larger coats seem to be constructed on similar principles, with the addition of further hexagons. From Pearse and Bretscher.[560]

complex oligomeric proteins bind to recognition or "sorting" sequences such as dileucine on YXXϕ (Y = Tyr, X = any amino acid, ϕ = bulky hydrophobic).[567a–d] The adapter proteins also bind to clathrin, the N-terminal β-propeller domain associating with the sequence LϕXϕD / E of some AP adapters.[567c] Completion of a coated vesicle requires membrane fusion as the vesicle is pinched off from the membrane surface. A GTPase (Chapter 11) called **dynamin** as well as another protein **endophilin I** are required. Endophilin I is an acyltransferase able to transfer the fatty acyl group of arachidonoyl-coenzyme A to lysophosphatidic acid in the cytosolic surface of the membrane. This may change the curvature of the membrane, assisting in vesicle formation.[567e]

Once inside a cell the vesicles lose their coats to become endosomes which may then fuse with lysosomes or with Golgi membranes. The removal of a clathrin coat requires ATP as well as the chaperonin Hsp 70 (Chapter 10) and a coat protein called **auxilin**.[568] Triskelion is distorted and displaced from the clathrin cage. The interior of the newly formed endosome is quickly acidified by the action of a proton pump in the vesicle walls.[554,569] This sometimes leads to dissociation of enclosed receptors from their ligands and permits recycling of receptors and lipids of the vesicle membranes to the cell surface. This is the case for the low-density lipoprotein receptor.[570,571]

**Exocytosis**, by which the content of a secretion vesicle is released to the outside of a cell, is just as important as endocytosis. The process is sometimes very specialized. For example, the release of a nematocyst from *Hydra* (Fig. 1-13) can occur in about 3 ms.[572] Exocytosis involves **fusion** of membranes,[562,573] a process also occurring during the movement of endosomes along the endocytic pathway,[574] during vesicular transport between Golgi compartments, and in many other biological processes. Exocytosis is often triggered by the binding of $Ca^{2+}$ to specific proteins of the vesicle wall and of the cytoskeleton (Chapter 7).[575] The fusion of membranes at several stages in the vesicle-mediated transport of materials between Golgi compartments requires a specific protein known as the *N*-ethylmaleimide-sensitive fusion protein (**NSF**).[576,577] A host of other specialized proteins are also involved[562,578] and are discussed in Chapter 10. See also Chapters 20 and 29.

## D. Communication

External coats and cell walls help to control the access of materials to a cell. However, it is the outer surface of the plasma membrane that makes the cell's first contact with nutrients, hormones, and other important chemicals. The membrane must often not only detect these materials but also send signals to the interior of the cell and sometimes to adjacent cells. These signals may be about changes in pH or nutrient concentration or the presence of hormones, neurotransmitters, or harmful materials. For these reasons cell membranes contain many embedded receptors and signaling complexes. These are discussed in Chapter 11 and later sections of the book.

The plasma membrane also contains many "markers" of the individuality of the species or of an individual. These are chemical groupings that, in higher animals, can be recognized by the immune system as "self" rather than as a foreign invader. These surface markers may also be used by parasites as camouflage to evade the immune response of the host. Such chemical groupings on cell surfaces are often described as **antigenic determinants** and the molecules that carry them as **antigens**. Each antigenic determinant elicits the production of antibodies that will bind specifically to it. Over 250 different antigenic groups have already been described for the surface of the red blood cell. They determine the blood type. Groups on the surfaces of other cells determine whether a transplanted tissue will be rejected. Various proteins from plant and other sources act as **agglutinins** by binding to surface groups much as do antibodies. Viruses that attack cells may also become adsorbed onto specific surface molecules, which act as receptors.

Immunoglobulins can also be receptors. For example, molecules of IgE bound to basophils and the related **mast cells** of tissues serve as receptors for allergens. Binding of an allergen to the IgE molecules stimulates the release of granules containing histamine and other substances (Chapter 31).

## E. The Extracellular Matrix and Cell Walls

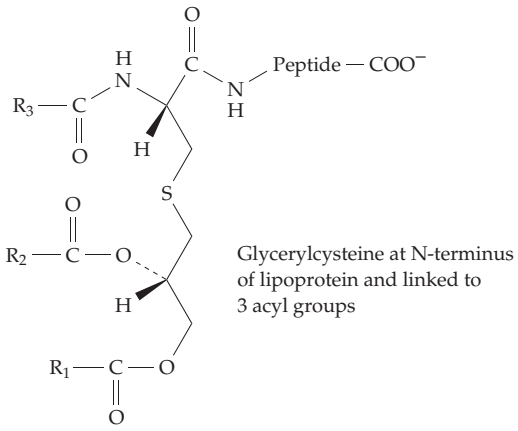
The surroundings of cells are extremely complex and vary from one organism to another and from one tissue to another. The principal function of cell walls and other surface coats is to protect cells against attack by organisms and against physical disruption.

### 1. The Structure of Bacterial Cell Walls

The plasma membrane of bacterial cells, other than the wall-less mycoplasmas and some archaebacteria, is surrounded by a multilayered wall which may be separated from the membrane by a thin **periplasm** (or periplasmic space). This can be seen most clearly in suitably prepared thin sections of cells of *E. coli* or other gram-negative bacteria as a relatively empty space of 11- to 25-nm thickness (Fig. 8-28).[579–581] The volume of this space (which may be filled with gelled material) depends upon the osmotic pressure of the medium. In *E. coli* it contains 20–40% of the total

cell water. In gram-negative bacteria the innermost layer of the walls lies within the periplasm (Fig. 8-28). It is a porous network of a highly crosslinked material known as **peptidoglycan** or **murein**. The backbone of the peptidoglycan is a β-1,4-linked polymer of alternating *N*-acetylglucosamine and *N*-acetylmuramic acid residues. Alternate units of the resulting chitin-like molecule carry unusual peptides attached to the lactyl groups of the *N*-acetylmuramic acid units (Fig. 8-29). The peptide side chains are crosslinked as indicated in the figure. The peptides vary considerably in structure and crosslinkages.[582,583] In *E. coli* and other gram-negative bacteria the peptidoglycan forms a thin (2-nm) continuous network around the cell, but in gram-positive bacteria the highly crosslinked peptidoglycan forms a layer as much as 10 nm thick.[584] The peptidoglycan layer is surrounded by other layers whose structures vary from one organism to another, with the outermost antigenic layers being the most variable.

***The outer membrane of gram-negative bacteria.*** Outside the murein layer of *E. coli* and other gram-negative organisms is a phospholipid-containing **outer membrane** which has the thickness and something of the structure of a typical biological membrane (Fig. 8-28).[585,586] This membrane is attached to the peptidoglycan layer with the aid of a small hydrophobic 58-residue **lipoprotein** whose N terminus contains a **glycerylcysteine** which carries three fatty acids connected by ester and amide linkages:




The fatty acid chains are evidently embedded in the outer membrane as an anchor. About one-third of the lipoprotein molecules are attached covalently to the peptidoglycan through an amide linkage between the side chain amino group of the C-terminal lysine of the protein and a diaminopimelic acid residue of the peptidoglycan (Fig. 8-29). Thus, the protein replaces one of the terminal D-alanine residues of about one in ten of the murein peptides. There are ~2.5 x $10^5$ molecules of the bound form of the lipoprotein per cell spread over a surface area of peptidoglycan of ~3 $\mu m^2$. They appear to be associated as trimers located primarily in the periplasmic space.[589]

About $10^5$ copies per cell of the previously mentioned (Section C,2) larger 325-residue structural protein, **OmpA protein**,[350] after its gene symbol *ompA*

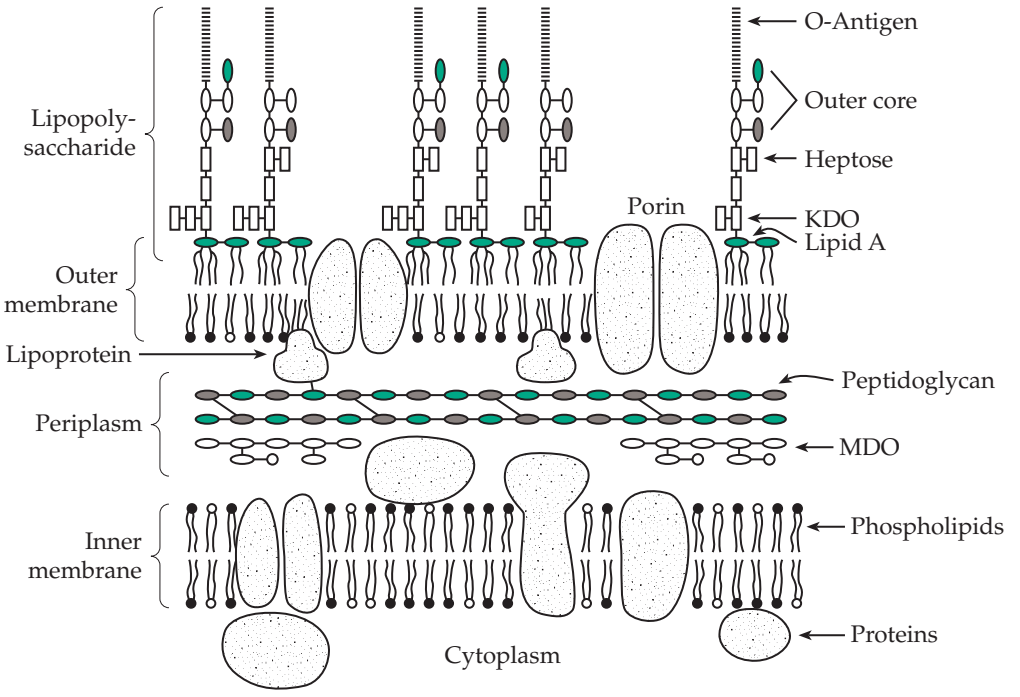
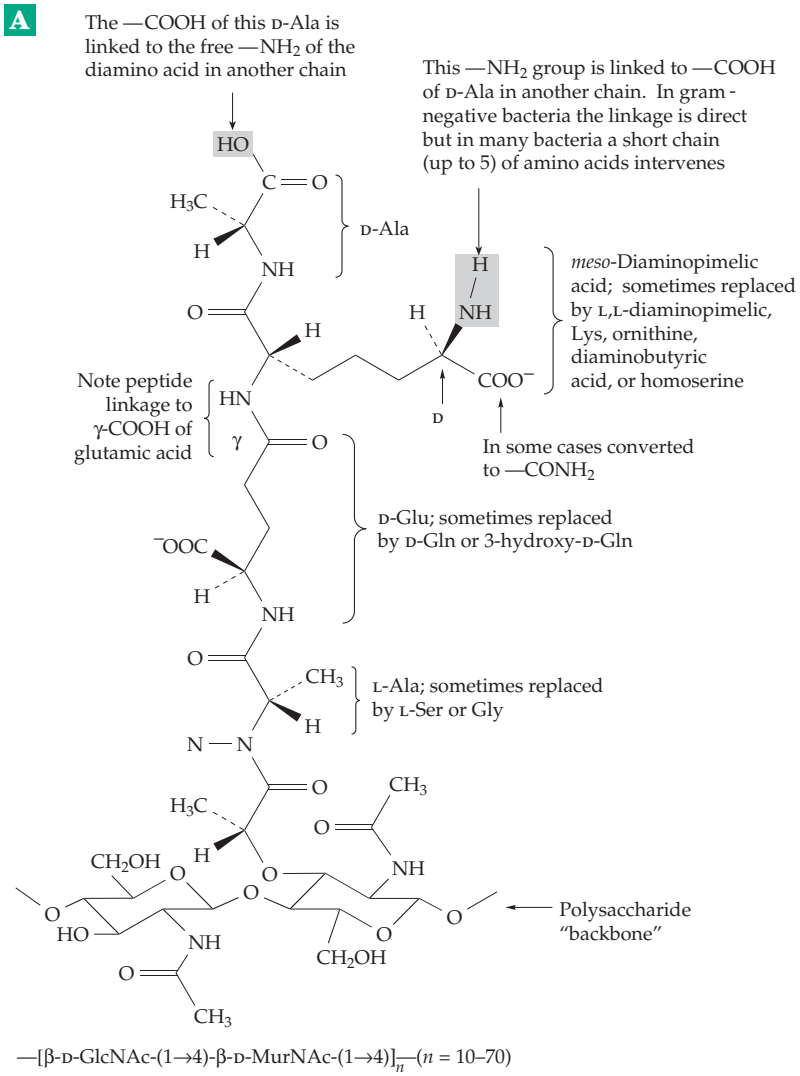



**Figure 8-28** Schematic molecular structure of the *E. coli* envelope. Sugar residues are represented by ovals and rectangles. Circles represent polar head groups of phospholipids. MDO, membrane-derived oligosaccharides; KDO, 3-deoxy-*manno*-octulosonic acid (structures for KDO and heptose are in Fig. 4-15). From Raetz and Dowhan.[587]




(outer membrane protein A), are also present. The outer membrane contains almost the same number of molecules of two porins. Together with phospholipids or with the lipopolysaccharide discussed in the next paragraphs, the porins and OmpA protein associate in hexagonal arrays which provide the basic framework structure of the outer membrane.[585] Two of the outer membrane proteins have been shown to contain some of the modified lysine α-aminoadipic acid 5-semialdehyde (**allysine**).[590] The aldehyde groups of allysine are able to form crosslinks to other proteins as has been well established for collagen and elastin in the human body.

A characteristic feature of the outer surface of gram-negative bacteria is a **lipopolysaccharide**[586] that is anchored in the outer membrane. It was discussed briefly in Chapter 4 where the structures of the repeating oligosaccharides known as O-antigens are given. Figures 8-28 and 8-30 show the manner in which the oligosaccharide bearing the O antigen is attached to a lipophilic anchoring group that is embedded in the outer membrane of the bacteria.[591] The anchor, which is called **lipid A**, is a β-1,6-linked disaccharide of *N*-acetylglucosamine. It is also linked both to phosphate groups and to the fatty acyl groups that fit into the lipid bilayer of the membrane. In *E. coli* and *Salmonella typhimurium* four molecules of **3-D-hydroxymyristic** acid are joined by ester and amide linkages to the two GlcN units (Fig. 8-30). Other fatty acids, including lauric, myristic, and palmitic acids, are esterified to the hydroxyl groups of two or three residues of hydroxymyristic acid (Fig. 8-30). Lipid A from other gram-negative bacteria is similar but with variations in the fatty acyl group composition and linkages to the carbohydrate.[592–594]

**Figure 8-29** (A) Repeating unit of structure of a bacterial peptidoglycan (murein). Some connecting bridges are pentaglycine (*Staphylococcus aureus*), trialanylthreonine (*Micrococcus roseum*), and polyserine (*S. epidermis*). (B) Schematic drawing of the peptidoglycan of *S. aureus*. From Osborn.[588]

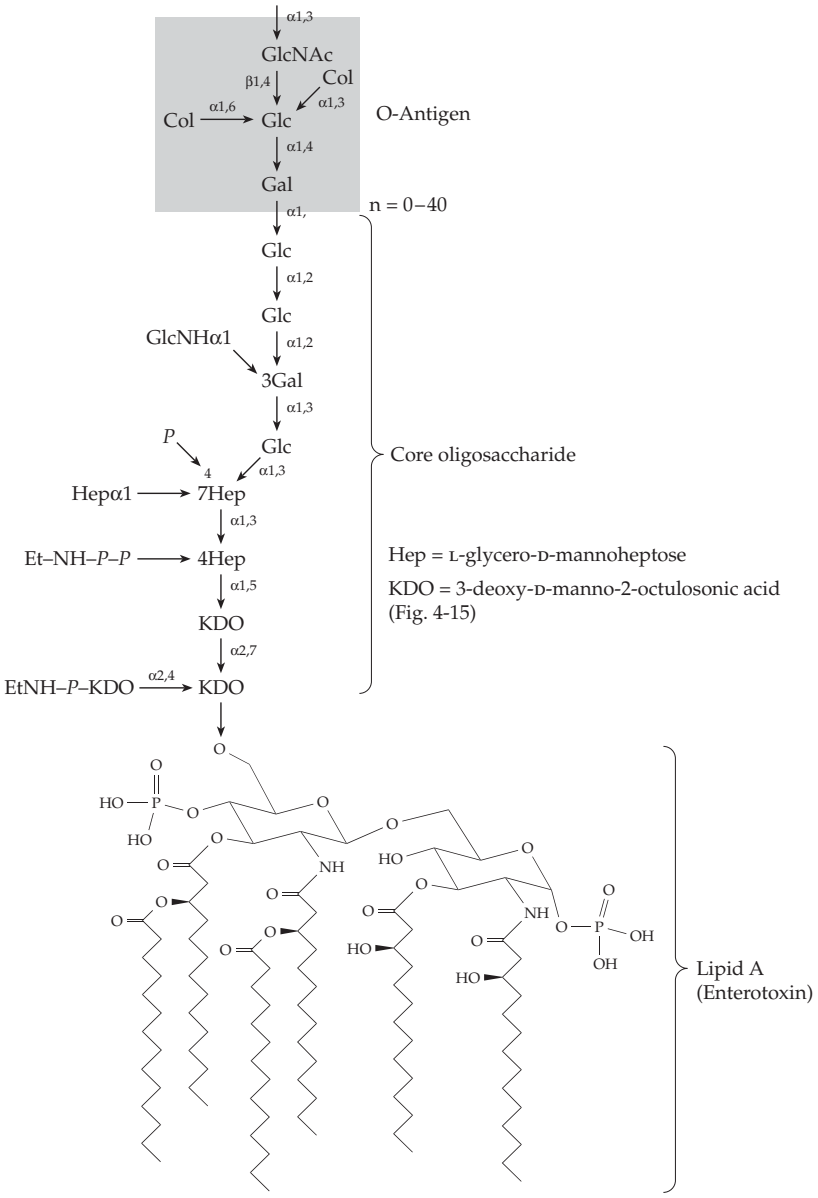
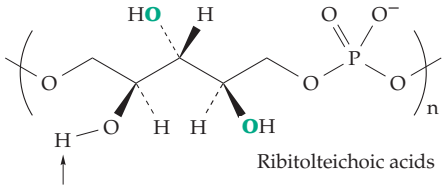


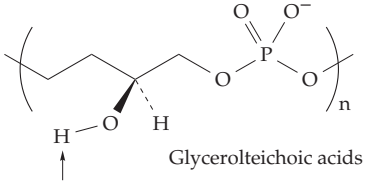

**Figure 8-30** Structures of the lipopolysaccharides of the outer membrane of *E. coli* and *S. typhimurium* including the bilayer anchor lipid A. For structures of L-glycero-D-mannoheptulose, KDO, and colitose, see Fig. 4-15.

In *Bacillus subtilus* D-alanine is attached to one of the green oxygen atoms in at least half of the units



Ribitolteichoic acids

In *B. subtilus* β-D-glucose is attached here



Glycerolteichoic acids

In *Lactobacillus arabinosus* D-alanine in ester linkage occupies this position but is replaced by β-D-glucose on about one unit in nine

Many differences are also found in the "core" oligosaccharide (Fig. 8-30) and in the O-antigens.[594–596]

Why do cells of *Salmonella* have a thousand distinguishable surface antigens, many based on differences in the O-antigens? The ends of these carbohydrate clusters are the groups to which the antibodies in animals clamp themselves if the bacteria enter the bloodstream. Mutants known as R forms (because of the growth as rough colonies on agar plates) completely lack the outer O-antigen. The R mutants of *Salmonella* are nonpathogenic, whereas the smooth strains with intact O-antigen often cause illness. Perhaps, if the O-antigen has the right cluster of sugar rings at the ends, the host organism does not recognize it as dangerous. This is only part of a continuous battle between the immune system of the body and camouflaged surfaces of attacking pathogens.

The lipopolysaccharides of bacterial surfaces have been identified as the "**endotoxins**" that cause many of the worst effects of gram-negative bacterial infections and that are often lethal.[597] Although the name endotoxin came from the assumption that there was an internal heat-resistant toxin that was released from the bacteria, the toxin is part of the bacterial cell wall. While the O-antigens provide for innumological recognition, it is the *lipid A* part that is responsible for the toxicity and unusually strong fever-inducing properties of the lipopolysaccharide.[592–594,597]

***Teichoic and teichuronic acids.*** The cell walls of gram-positive bacteria are composed of a thick peptidoglycan layer which also contains proteins and additional polymers known as **teichoic acids** and **teichuronic acids**. In some species these account for 50% of the dry weight of the cells.[598,599] Teichoic acids are high polymers of the following general types: They are often attached through phosphodiester linkages to *N*-acetylglucosamine or a disaccharide which, in turn, is attached to muramic acid residues of the peptidoglycan. Since the teichoic acid is uniformly distributed in its attachment to the peptidoglycan, it must be intimately associated with peptidoglycan throughout the wall.

Teichoic acids are often covalently attached to glycolipids which are part of the plasma membrane. For example, the glycerolteichoic acid of *Streptococcus faecium* contains about 28 monomer units of glycerol phosphate, approximately 60% of which carry residues of kojibiose (Glu$\alpha$1 $\rightarrow$ 2Glu) as a phosphatidylkojibiosyl diacylglycerol membrane anchor.[600] Teichuronic acids contain uronic acids:

$$[\rightarrow 4(N\text{-acetyl})\text{–ManA}\beta 1\rightarrow 6\text{Glc}\alpha 1\text{–}]_n$$

From *Micrococcus luteus*[601]

$$\left[\rightarrow 4\text{Glc}\beta 1\rightarrow 3\text{Rha}\alpha 1\rightarrow 4\text{Rha}\alpha 1\text{–}\begin{array}{c}3\\\uparrow\\\text{GlcA}\beta 1\end{array}\right]_n$$

From *Bacillus megaterium*[602]

Nearly 800 of the foregoing tetrasaccharide units are joined to give a large, densely packed, ~500-kDa polymer. The reducing end of the polymer is covalently attached to the peptidoglycan layer.

***Mycobacteria.*** Most mycobacterial species seem to be harmless saprophytic soil bacteria of the gram-positive group. However, tuberculosis (caused by *Mycobacterium tuberculosis*) may infect one-third of the inhabitants of the earth[603] and kills about three million people per year.[604] Leprosy, caused by *M. leprae*, affects 10 million or more.[605] Most antibiotics are not effective against mycobacteria because of the unusual nature of their multilayered cell walls. The muramic acid of their peptidoglycans contains glycolyl groups instead of acetyl groups, perhaps providing extra hydrogen bonding within the wall.[604] Most characteristically, they are rich in lipid materials. Among these are a unique polysaccharide, a highly branched **arabinogalactan**, that is covalently attached to muramic acid residues of the peptidoglycan. Clusters of arabinofuranosyl units at the nonreducing ends of the chains are esterified with **mycolic acids** (Section A,1).[10,604,605] Mycobacteria also contain **lipoarabinomannans** that act as major antigens.[603,606] Some mycobacterial species synthesize small glycopeptidolipids that may disrupt cell membranes of hosts.[607] Mycobacteria don't produce toxins of usual types but cause slow, long-lasting infections.

***Other bacterial coats.*** Archaebacteria not only have unusual plasma membranes that contain phytanyl and diphytanyl groups (Section A,3)[608] but also have special surface layers (**S-layers**) that may consist of many copies of a single protein that is anchored in the cell membrane.[609] The surface protein of the hypothermic *Staphylothermus marius* consists of a complex structure formed from a tetramer of 92-kDa rods with an equal number of 85-kDa "arms."[610,611] S-layers are often formed not only by archaebacteria but also by eubacteria of several types and with quite varied structures.[612–614] While many bacteria carry adhesins on pili, in others these adhesive proteins are also components of surface layers.[615] Additional sheaths, capsules, or slime layers, often composed of dextrans (Chapter 4) and other carbohydrates, surround some bacteria.

## 2. The Surroundings of Animal Cells

Cells in the external epithelial layers are always surrounded by a protective covering of some kind. Our own skin is made up of specialized cells which become filled with microfibrils of keratin as they move outward and become the relatively dry nonliving external surface (Box 8-F). Internal epithelial cells secrete protein and carbohydrate materials that form a thin **basement membrane** around the exposed parts of the cells. The **connective tissue** that lies between organs and which also includes tendons, cartilage, and bone consists of a relatively small number of cells surrounded by a "matrix" consisting of the protein fibers collagen and elastin in a "ground substance" rich in proteoglycans (Chapter 4).[616–618] In bone, the calcium phosphate is deposited within this matrix.

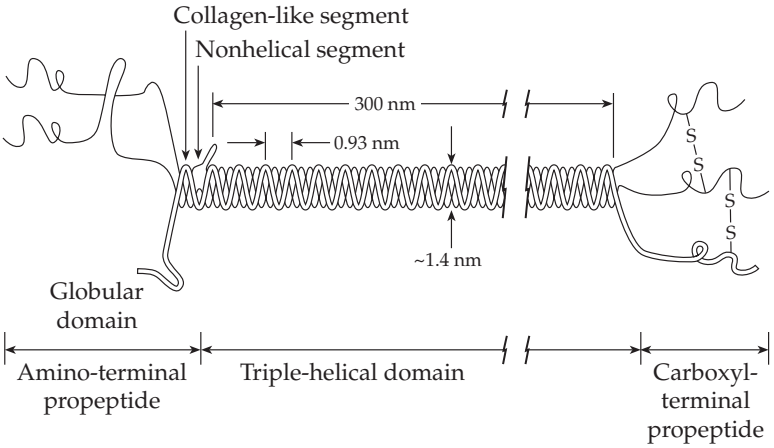
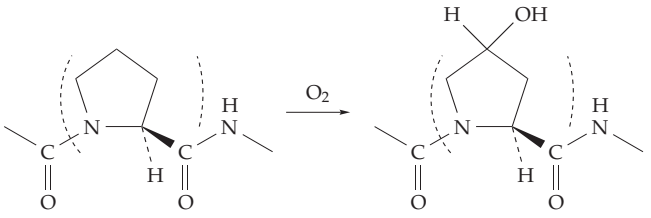



**Figure 8-31** Schematic representation of type I procollagen. The molecule is composed of two identical proα1 chains (solid lines) and one proα2 chain (dashed line). In addition to the central triple-helical region that gives rise to the collagen molecule, as portrayed in Fig. 2-23(C), the precursor contains amino- and carboxyl-terminal non-triple-helical domains. The amino-terminal domain is composed of a presumably globular region, a short collagen-like segment, and a non-triple-helical region in which cleavages by the amino-terminal protease occur. Interchain disulfide bonds are limited to the carboxyl-terminal domain. The short telopeptides at the ends of collagen α chains represent the residual sequences of the linkage regions between the collagen helix and the terminal domains. After Prockop[628] and Byers.[625]

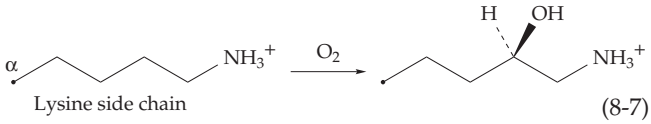
***The collagens.*** The most abundant proteins in the body are the collagens,[619–624] a family of closely related materials that account for 20% of the total protein in higher animals. Collagens make up much of the organic mass of skin, tendons, blood vessels, bone, the cornea, vitreous humor of the eye, and basement membranes. They are found in every metazoan phylum studied. There are at least 16 types in the human body.[625] Type I collagen, which accounts for 90% of the total, is the major form occurring in skin, tendon, and bone. It is synthesized by the fibroblasts and is excreted into the extracellular space where it is polymerized into a durable long-lived material. Collagen II is found exclusively in cartilage and the vitreous humor of the eyes. Form III is located in blood vessels and intestines and is prominent in embryonic tissues. Collagen IV is the major form found in basement membranes. Some other collagens and characteristic features are listed in Table 8-4.

All collagens contain the triple-helical structure shown in Fig. 2-23. Collagen I consists of two chains of one kind (α1) and one of another (α2), while most other collagens have three identical chains. In these chains over 1000 residues have the characteristic GlyXY sequence.[624,626] At each end of the rodlike molecules short segments of the peptides fold into globular domains. For type I collagen 16 residues at the N termini and 25 at the C termini form these domains. Collagens are synthesized as intracellular precursors known as **procollagens**. The three chains of procollagens are much longer than in the mature proteins and have larger non-triple-helical ends (Fig. 8-31). The C-terminal extensions are crosslinked by S–S bridges.[625,627] Synthesis of the collagen chains requires at least six minutes after which they are released into the cisternal space of the ER, associate, and become crosslinked (Fig. 8-32). This crosslinking ensures that the three chains remain in proper register in the triple helix while the procollagen is converted into collagen, a process involving additional crosslinking.

Before this "maturation" can occur there must be other modifications to procollagen. These begin while the peptide chains are still attached to ribosomes of the rough ER. Hydroxylases (Chapter 18) localized in the membranous vesicles of the ER convert some of the proline and lysine residues of the procollagen chains into **4-hydroxyproline**[625,629,630] and **hydroxylysine** (Eqs. 8-6 and 8-7). Lesser amounts of 3-hydroxyproline are formed. About 100 molecules of 4-hydroxyproline and 50 of 5-hydroxylysine are created in each α1 chain.




(8-6)




(8-7)

Galactosyl units are then transferred onto some of the hydroxyl groups of the hydroxylysine side chains, and glucosyl groups are transferred onto some of the galactosyl groups. This glycosylation may prevent incorrect association of the procollagen molecules.

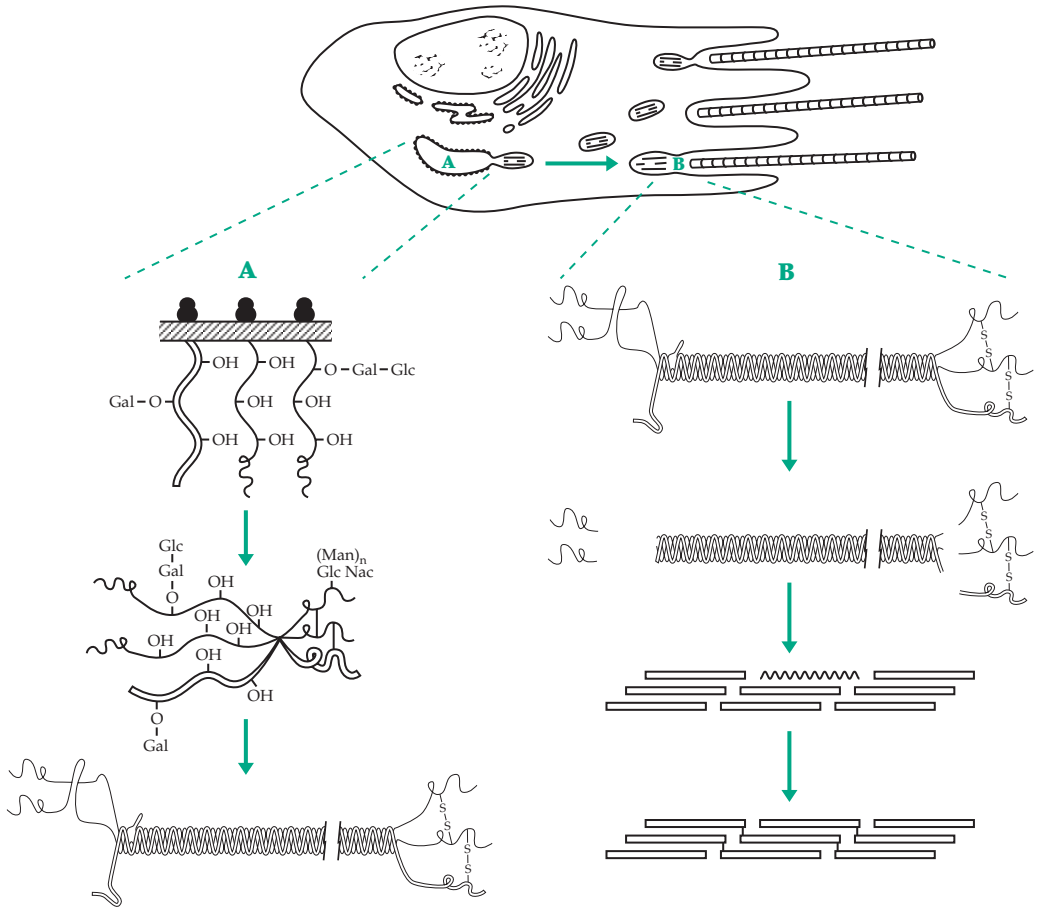


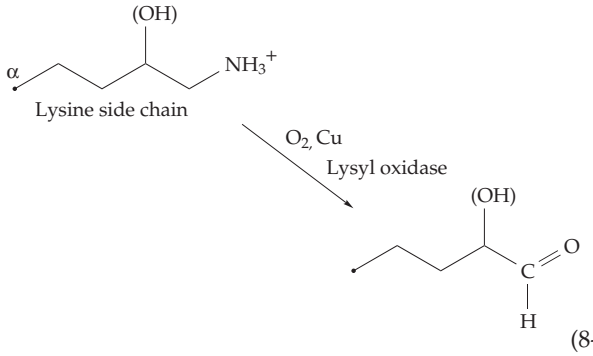

**Figure 8-32** Scheme summarizing the biosynthesis of a fibrillar collagen by a fibroblast. (A) Assembly of pro-α chains in cisternae of the rough endoplasmic reticulum; posttranslational hydroxylations and glycosylations; association and disulfide bonding of C-propeptides; and zipper-like folding of the triple helix by nucleated growth. (B) Proteolytic processing of procollagen to collagen; self-assembly of fibrils by nucleated growth; and covalent crosslinking of the fibrils. The collagen molecule is first shown as a triple helix and then either as a *wavy line* to depict a molecule assembling on the surface of a fibril or as a *rectangle* to depict the quarter-staggered assay of monomers in a fibril. The proteolytic processing of procollagen and assembly of fibrils may occur within crypts of fibrils as shown here or perhaps at some distance from the cell. After D. J. Prockop.[628]

Within the extracellular space two **procollagen peptidases** act to cleave a 35-kDa peptide from the C terminus[631] and a 20-kDa peptide from the N-terminal end of each of the three chains of the secreted procollagen. The amino acid composition of the peptides removed is quite unlike that of the remaining collagen monomer (also called tropocollagen) which contains one-third glycine and much proline.

The three-stranded monomers of collagens I–III are rods of dimensions ~1.4 x 300 nm (Fig. 8-23). When they reach their final location they associate and become crosslinked to form strong fibers with diameters ranging from 8 nm to 0.5 m. Tendons tend to contain large fibrils, while those in bone are small. The smallest 8-nm fibrils must contain about 20 triple helices in a single cross section but the successive monomers are staggered by 6.4–6.7 nm (234 residues for type I collagen) in such a way that 3-nm gaps are left between the ends (Fig. 8-32). These gaps give rise to the characteristic banding pattern seen in the fibrils in Fig. 2-23D. Finer analysis of the bands together with the known sequences shows that the bands also reflect the locations of residues with charged side chains.[632] These ionized groups are thought to provide electrostatic stabilization to the fibrils.[633] The exact packing of the collagen molecules into sheetlike or microfibrillar crystalline arrays is still uncertain.[634] However, there is agreement on the staggered arrangement[634a] and upon the fact that precisely formed crosslinks with neighboring rods prevent the gaps from weakening the fibrils.

Crosslinking of collagen is initiated by oxidation of some of the lysyl and hydroxylysyl side chains from amino groups to aldehyde groups under the action of a copper-containing oxidase (Eq. 8-8, Chapter 18). The aldehyde groups enter into a variety of reactions that

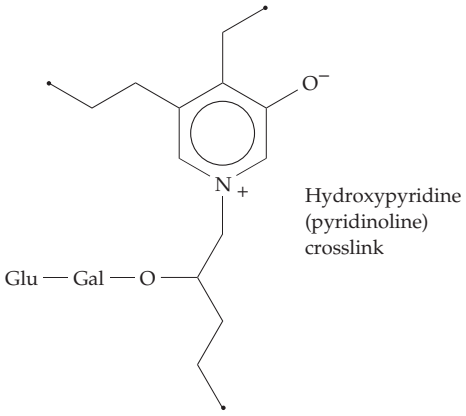
lead to crosslinking of the collagen monomers and to

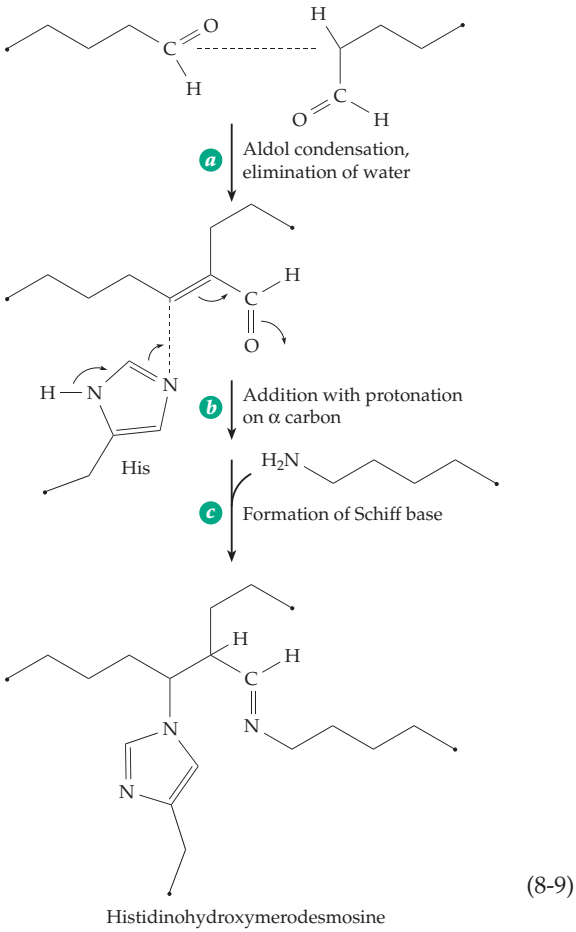



(8-8)

the formation of insoluble fibers.[635] One reaction is an aldol condensation followed by elimination of water (Eq. 8-9, step *a*). If one of the two aldehydes involved in the condensation is derived from hydroxylysine and the other from lysine, two isomeric condensation products are formed. The aldol condensation product can react further: An imidazole group from a histidine side chain can add to the carbon–carbon double bond and (either before or after this reaction) another lysine side chain can form a Schiff base with the free aldehyde. The results of these two processes are summarized in Eq. 8-9, step *b*. The final product **histidinohydroxymerodesmosine** links four different side chain groups. In other instances, simple Schiff bases (**aldimines**) are formed between aldehyde and ε-amino groups. If there is an adjacent hydroxyl group these can isomerize to ketoamines (Eq. 8-10). Two residues of hydroxylysine, one of which is glycosylated as shown in Eq. 8-10, are often involved. Borohydride reduction (Eq. 4-2) and hydrolysis leads to isolation of **dihydroxylysinonorleucine**, the predominant product of such treatment of bone or cartilage.

Crosslinkages reducible with borohydride are characteristic of newly formed collagen but these disappear and are replaced by more stable crosslinks as collagen matures. For example, a **3-hydroxypyridine** that joins three triple helices may be formed from the reaction of two ketoamine groupings (with elimination of one glycosylated hydroxylysine residue).[636–639] Similar chemistry can also produce pyrrole crosslinks.[639]




(8-9)

Histidinohydroxymerodesmosine

The crosslinkages in collagen are not located at random but are found in certain positions, often toward the ends of the collagen monomers. Thus, histidine residues are found only at positions 89, 929, and 1034 in the α1(I) peptide chains. Residue 9 in the N-terminal globular portion of one chain is linked to residue 946 of another while residue 103 is linked to 1047 in the globular C-terminal peptide.[640] The variety and number of crosslinkages vary among different species. As collagen ages through a lifetime, glycation (Eq. 4-8) leads to crosslinkages in which two ketoamines are formed via glycation cyclize.[641]

Collagens I, II, and III form fibrils with similar structures. However, other collagens are longer or shorter and aggregate in different ways. A pepsin-resistant part of the collagen V molecules resembles collagens I, II, and III but it may contain an additional non-collagenous segment at the N terminus. Collagens V and XI are quantitatively minor components of the extracellular matrix but are thought to provide a core around which the fibrils of collagens I and III may form.[642] Types XII, XIV, IX, and XVI collagens contain interruptions in the helix which create bends, flexible sites, and sites of increased proteolytic susceptibility. They may link the fibrils to other components of the surrounding matrix.[643]

## TABLE 8-4
## Types of Vertebrate Collagen

| Type number | Location | Characteristics | Gene location: human chromosome |
|---|---|---|---|
| **Forming quarter-staggered fibrils** | | | |
| I[a,b] | Skin, bone, tendon, dentin | Most abundant, banded quarter-staggered fibrils | 7, 17 |
| III[a,c] | Skin, blood vessels | Abundant, small banded fibrils | 2 |
| V[a,d] | Most interstitial tissues; cartilage, bone | Abundant, small fibrils | 2 |
| **Predominant in cartilage and bone** | | | |
| II[a] | Hyaline cartilage, vitreous humor | Very abundant, small banded fibrils | 12 |
| XI[a] | Hyaline cartilage | | |
| **With interrupted triple helices** | | | |
| XII[a,e] | Embryonic tendon, periodental ligaments | Fibril associated | |
| XIV[a,e] | Fetal skin, tendon | Fibril associated | |
| IX[a,e] | Cartilage, vitreous humor | Minor, contains attached glycosamino-glycan, fibril associated | |
| XVI[f] | Cartilage | Fibril associated | |
| **Forming sheets and networks** | | | |
| IV[a,g] | All basement membranes | Nonfibrillar network | 13 |
| X[a] | Mineralizing cartilage, growth plate | Short chain | |
| VIII[a,h] | Endothelial cells; Descemet's membrane of the cornea | Small helices linked in hexagonal arrays | |
| **Forming beaded filaments** | | | |
| VI[a] | Most interstitial tissues, intervertebral discs | Beaded microfilaments | |
| **Forming anchoring fibrils** | | | |
| VII[a,i] | Basement membranes | Long-chain, antiparallel dimers, anchoring fibrils | |

[a] Martin, G. R., Timpl, R., Müller, P. K., and Kühn, K. (1985) *Trends Biochem. Sci.* **10**, 285-287
van der Rest, M., and Garrone, R. (1991) *FASEB J.* **5**, 2814-2823
Prockop, D. J., and Kivirikko, K. I. (1995) *Ann. Rev. Biochem.* **64**, 403-434

[b] Prockop, D. J. (1990) *J. Biol. Chem.* **265**, 15349-15352

[c] Nah, H.-D., Niu, Z., and Adams, S. L. (1994) *J. Biol. Chem.* **269**, 16443-16448

[d] Myers, J. C., Loidl, H. R., Stolle, C. A., and Seyer, J. M. (1985) *J. Biol. Chem.* **260**, 5533-5541

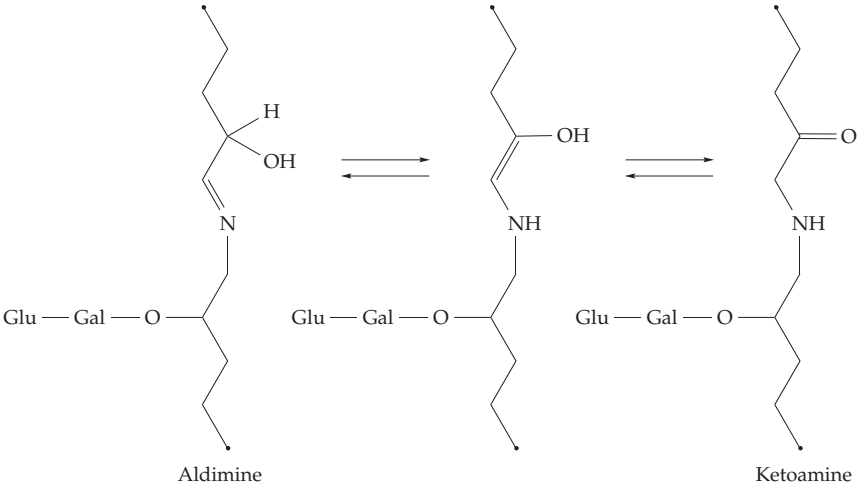
[e] Shaw, L. M., and Olsen, B. R. (1991) *Trends Biochem. Sci.* **16**, 191-194

[f] Myers, J. C., Yang, H., D'Ippolito, J. A., Presente, A., Miller, M. K., and Dion, A. S. (1994) *J. Biol. Chem.* **269**, 18549-18557

[g] Hudson, B. G., Reeders, S. T., and Tryggvason, K. (1993) *J. Biol. Chem.* **268**, 26033-26036

[h] Benya, P. D., and Radilla, S. R. (1986) *J. Biol. Chem.* **261**, 4160-4169

[i] Lunstrum, G. P., Sakai, L. Y., Keene, D. R., Morris, N. P., and Burgeson, R. E. (1986) *J. Biol. Chem.* **261**, 9042-9048




Aldimine Ketoamine

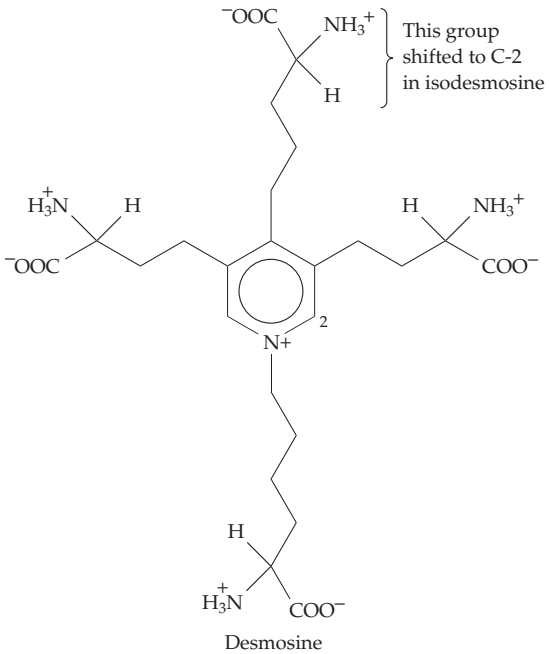
(8-10)

The basement membrane collagen IV forms a nonfibrillar network.[644,645] The 400-nm-long molecules aggregate via their identical ends. Four molecules are held together through their triple-helical N termini, while the C-terminal globular domains connect pairs of molecules. Type IV collagen is also a proteoglycan with a glycoaminoglycan chain attached to one of its nonhelical domains. It may become covalently attached to type II collagen via a hydroxypyridine linkage.[646] Dimers of the microfibrillar collagen VI are formed from 105-nm-long monomers by antiparallel and staggered alignment, with the 75-nm overlapping helical segments twisting around each other to form coiled dimers. After a symmetrical association of dimers to tetramers, fibrillar structures are formed by end-to-end aggregation. The connection of monomers to dimers, tetramers, and polymers occurs by disulfide bonds between triple-helical segments and globular domains. The 450-nm-long collagen VII molecules associate into antiparallel dimeric structures which show a 60-nm overlap and which subsequently assemble laterally with their ends in register.[622]

At least 32 genes encode the α peptide chains of vertebrate collagens.[647,648] These chains are assembled into the 19 known types of collagen. Alternative splicing of the mRNAs provides additional isoforms.[649] The collagen α2(I) gene from both the chick and human DNA is ~38 kb in length and consists of 52 exons separated by introns ranging in length from 80 to 2000 base pairs.[647,650] At least nine of the exons that encode the triple-helical regions have exactly 54 bp. All are multiples of 9 bp, i.e., the length needed to encode one Gly-X-Y triplet (Chapter 2, Section D,4). The significance of these observations is unclear but the presence of so many introns does suggest ways in which collagen sequences could have been transferred into the genes for such proteins as acetylcholineesterase and the C1q component of complement.[647] The human α1(I) collagen gene also consists of 51 segments but they lie within a shorter 18-kb length of DNA.[651] A collagen gene from *Drosophila* is much less fragmented.

Collagens are found in all metazoan organisms.[652] Invertebrate collagens play a variety of specialized roles.[653] For example, minicollagens from *Hydra* strengthen the walls of their nematocysts.[572]

***Elastic fibers.*** The elastic properties of lung, skin, and large blood vessels are provided by elastic fibers in the extracellular matrix.[654] The fibers consist of amorphous material together with the insoluble protein **elastin**, which is rich in glycine, proline, and hydrophobic amino acids. Its special structure (Chapter 2) provides elasticity to the fibers. A 72-kDa precursor **tropoelastin** is secreted into the extracellular space where it is acted upon by lysyl oxidase (Eq. 8-7) and crosslinked into a rubber-like network.[654] Remarkable crosslinkages are formed. Three aldehyde groups derived by oxidation of lysine side chains combine with one lysine amino group through aldol condensations, dehydration, and oxidation to form residues of **desmosine** and **isodesmosine**.




Desmosine

***Cartilage and basement membranes.*** Tendons consist largely of collagen, but in most tissues the collagen fibrils are embedded in a matrix of proteoglycans and various other proteins.[655–657] Both the core proteins of the proteoglycans and the attached

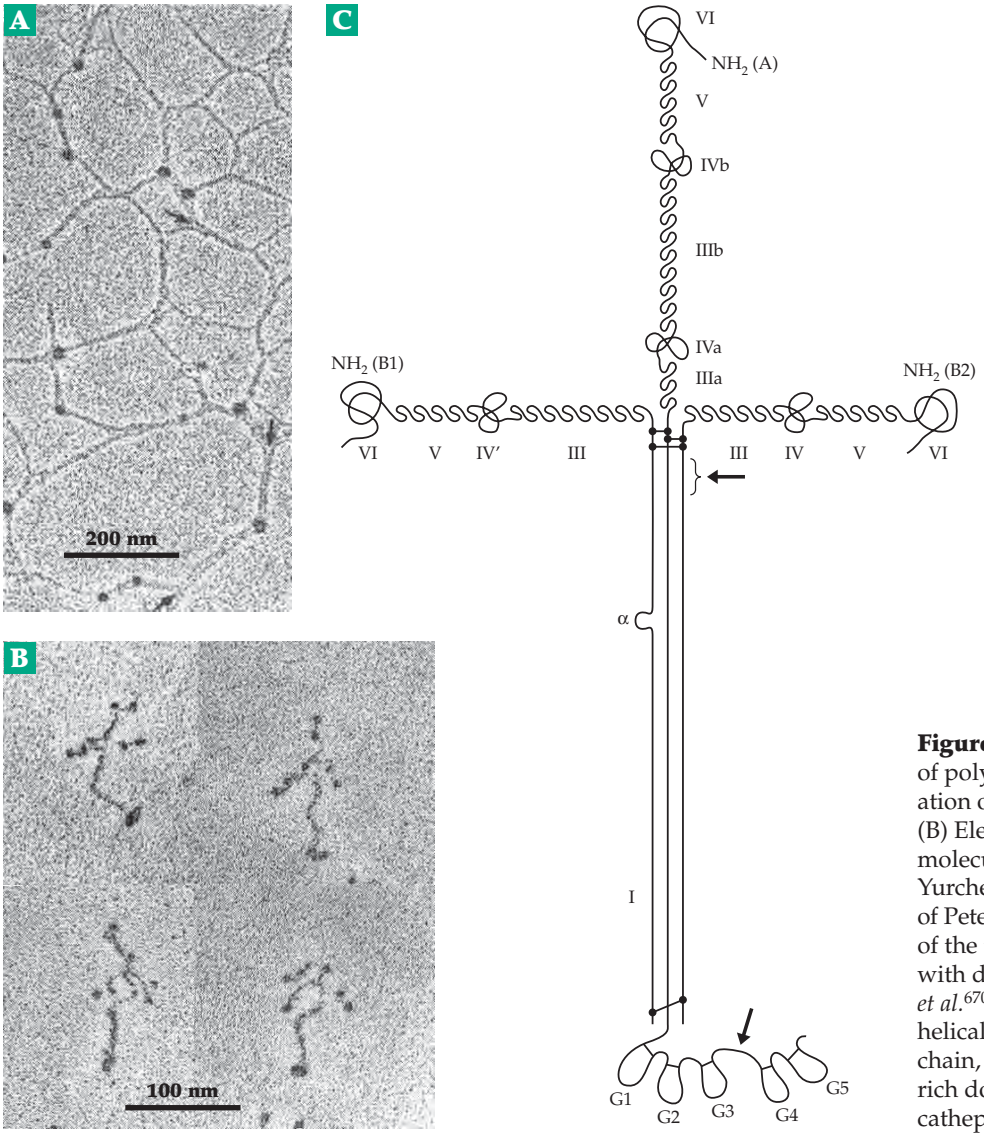



**Figure 8-33** (A) Electron micrograph of polygonal network formed by association of collagen type IV monomers. (B) Electron micrographs of single molecules of laminin. (A) and (B) from Yurchenko and Schittny.[663,663a] Courtesy of Peter Yurchenko (C) Structural model of the three-chain laminin molecule with domain designations. From Beck *et al.*[670] Domains I and II are a triple-helical coiled coil rod which, in the B1 chain, is interrupted by a small cysteine-rich domain α. Sites of cleavage by cathepsin are marked by arrows.

polysaccharides interact with collagen fibrils and with fibronectin and other (previously discussed) cell surface proteins. The cartilage matrix[658,659] consists largely of proteoglycans (Fig. 4-16)[660] and of several difficult to study, insoluble proteins. One of these is the 148-kDa **cartilage matrix protein**, which yields 52-kDa subunits upon reduction.[661,662] It interacts with both proteoglycans and collagen and may help to integrate the cartilage matrix.

Basement membranes (Fig. 1-6)[663] function in part as an exoskeleton that helps keep cells positioned. However, the thick basement membranes of the capillary walls of the glomeruli of the kidney provide the ultrafilters that prevent most proteins from entering the urine. Basement membranes contain large amounts of collagen IV, which forms a polygonal network (Fig. 8-33A). A second macromolecular network is formed by the very large 950-kDa cross-shaped multisubunit protein called **laminin** (Fig. 8-33B).[664–666] Laminin is one of a series of extracellular proteins which appear to have arisen by shuffling of structural modules during evolution.[667] It contains sites for binding to heparin, to integrins,[668] to the heparin sulfate proteoglycan **agrin** (see also Fig. 4-11), and to the 150-kDa sulfated glycoprotein **nidogen** (entactin).[669] At least seven isoforms of laminin, with varying tissue distributions, are formed, in part as a result of alternative splicing of the mRNA transcripts.[666] Laminin is rich in EGF-like modules.[665] An X-ray structure of three of them shows that they form a continuous rod of complex structure.[321] A smaller 100 kDa basement membrane **fibulin** also contains multiple EGF-like repeats.[671] As with other extracellular structures, crosslinking of laminin and other components of basement membranes by transglutaminase provides additional stability.[672]

## BOX 8-E GENETIC DEFECTS OF COLLAGEN STRUCTURE

Any major protein of the body is likely to be associated with a number of genetic problems. In the case of collagen, the possibility for harmful mutations is enhanced by the existence of a large number of genes that encode the more than 16 types of collagen which are expressed differently in different tissues. There are known human disorders resulting from defects in synthesis, secretion, or structure of types I, II, III, IV, and VII collagens. Other defects involve lysyl hydroxylase and procollagen *N*-proteinase.[a–d]

In the severe lethal form of **osteogenesis imperfecta** (brittle bone syndrome) the victims' collagen I may contain an α1 chain lacking as many as 100 residues or a shortened α2 chain. In other cases a cysteine, arginine, or other amino acid has been substituted for glycine in the triple-helical region of an α1 chain.[e,f] The cloning of collagen genes has permitted an exact description and precise location of the defects in these genes. Alterations toward the N-terminus of the α1 chain or in the α2 chain often cause a milder type of osteogenesis imperfecta.[g] Some patients have deletions in the pro-α2 chains of collagen I causing the chain to be incorporated into the collagen without removal of the N-terminal or C-terminal peptide to give a protein with poor stability.[h] In other cases amino acid substitutions in the α2 chain cause formation of chains with excessive posttranslational modification. Sometimes the α2 chain is not incorporated into the triple helix and the collagen I formed contains three α1 chains.

Another well established abnormality of collagen is found in cattle suffering from **dermatosparaxis**, a disease in which the skin is extremely brittle. The collagen chains are disorganized and have poor fiber-forming properties. The procollagen peptidase that cleaves a peptide from the N termini of the chains of procollagen is apparently defective. A similar human collagen disease is the **Ehlers–Danlos syndrome**, which in some instances is accompanied with recurrent joint dislocations and curvature of the spine. At least ten different types of the disorder are known.[d] The procollagen peptidase is sometimes lacking.[i] In other cases a person synthesizes an abnormal pro-α2 chain that is resistant to cleavage by the peptidase because of deletion of the normal cleavage site. In others collagen is formed in only small amounts or is degraded rapidly. Some individuals lack lysyl hydroxylase and others have a defect in mRNA splicing which causes loss of an exon from the mRNA and synthesis of shortened pro-α2 chains.[j]

Somewhat similar symptoms are observed in **lathyrism**, a disease which arises when animals ingest seeds of *Lathyris odoratus*, the common sweetpea. Since lathyris peas form part of the diet of some peoples, the condition is also known in humans and often causes curvature of the spine and rupture of the aorta. The biochemical problem has been traced to the presence in the seeds of **β-cyanoalanine** and of its decarboxylation product **β-aminopropionitrile**.

$$N \equiv C - CH_2CH_2 - NH_3^+$$

Although the mode of action is not certain, this compound is an inhibitor of lysyl oxidase essential to the crosslinking of both collagen and elastin. A hereditary defect with a similar effect in the mouse involves a defect in lysyl oxidase.[k,l]

Collagen defects account for a variety of other skeletal problems[m] including some cases of the common **osteoarthritis**.[n] Mice lacking the α1 chain of collagen IX develop a degenerative joint disease resembling human osteoarthritis.[o] An inherited defect in the basement membrane collagen IV is responsible for the inherited **Alport disease** in which kidney filtration is defective.[p,q] Similar symptoms are observed with the acute autoimmune disease **Goodpasture syndrome** and in **diabetic nephropathy**.[p]

---


[a] Prockop, D. J. (1990) *J. Biol. Chem.* **265**, 15349–15352

[b] Kuivaniemi, H., Tromp, G., and Prockop, D. J. (1991) *FASEB J.* **5**, 2052–2060

[c] Prockop, D. J., and Kivirikko, K. I. (1995) *Ann. Rev. Biochem.* **64**, 403–434

[d] Byers, P. H. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4029–4077, McGraw-Hill, New York

[e] Cohen-Solal, L., Zylberberg, L., Sangalli, A., Gomez Lira, M., and Mottes, M. (1994) *J. Biol. Chem.* **269**, 14751–14758

[f] Lightfoot, S. J., Atkinson, M. S., Murphy, G., Byers, P. H., and Kadler, K. E. (1994) *J. Biol. Chem.* **269**, 30352–30357

[g] Marini, J. C., Lewis, M. B., Wang, Q., Chen, K. J., and Orrison, B. M. (1993) *J. Biol. Chem.* **268**, 2667–2673

[h] Mundlos, S., Chan, D., Weng, Y. M., Sillence, D. O., Cole, W. G., and Bateman, J. F. (1996) *J. Biol. Chem.* **271**, 21068–21074

[i] Holmes, D. F., Watson, R. B., Steinmann, B., and Kadler, K. E. (1993) *J. Biol. Chem.* **268**, 15758–15765

[j] Weil, D., D'Alessio, M., Ramirez, F., Steinmann, B., Wirtz, M. K., Glanville, R. W., and Hollister, D. W. (1989) *J. Biol. Chem.* **264**, 16804–16809

[k] Pope, F. M., Martin, G. R., Lichtenstein, J. R., Penttinen, R., Gerson, B., Rowe, D. W., and McKusick, V. A. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1314–1316

[l] Rowe, D. W., McGoodwin, E. B., Martin, G. R., and Grahn, D. (1977) *J. Biol. Chem.* **252**, 939–942

[m] Freisinger, P., Ala-Kokko, L., LeGuellec, D., Franc, S., Bouvier, R., Ritvaniemi, P., Prockop, D. J., and Bonaventure, J. (1994) *J. Biol. Chem.* **269**, 13663–13669

[n] Ala-Kokko, L., Baldwin, C. T., Moskowitz, R. W., and Prockop, D. J. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6565–6568

[o] Fässler, R., Schnegelsberg, P. N. J., Dausman, J., Shinya, T., Muragaki, Y., McCarthy, M. T., Olsen, B. R., and Jaenisch, R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5070–5074

[p] Zhou, J., Hertz, J. M., Leinonen, A., and Tryggvason, K. (1992) *J. Biol. Chem.* **267**, 12475–12481

[q] Gunwar, S., Ballester, F., Noelken, M. E., Sado, Y., Ninomiya, Y., and Hudson, B. G. (1998) *J. Biol. Chem.* **273**, 8767–8775

## BOX 8-F SKIN

Mammalian skin must be tough, water-resistant, self-renewing, and rapidly healing. The outer layers of cells or **epidermis** consist principally of **keratinocytes**, epithelial cells specialized for formation of keratin (Fig. 7-31). In the inner layer of the epidermis the **basal stem cells** divide, providing a constant outward flow of cells which become progressively flattened, dehydrated, and filled with keratin fibrils.[a] The outer layers contain only dead cells which are finally sloughed or abraded from the surface. Human epidermis is completely renewed in about 28 days!

About 25 different human genes encode the keratins of skin and other soft tissues. Others specify the keratins of hair and nails.[b] Both of these hard tissues as well as claws, hoofs, beaks, horns, scales, quills, and feathers are largely keratin. However, there are additional constituents. During the final stages of keratinocyte differentiation a 15-nm-thick crosslinked sheath of protein, the **cornified cell envelope** (CE), forms beneath the plasma membrane.[c] Crosslinkages between keratin and other proteins are formed by the action of transglutaminases.[d–f] A specialized protein **involucrin**, which contains glutamine-rich repeating sequences, provides many of the side chain amide groups for the crosslinking reaction (Eq. 2-23).[g] **Loricin**, a protein containing glycine-rich flexible loops,[h] is also a major partner in these crosslinkages.[c,h] The histidine-rich **filaggrin** undergoes complex processing before binding to keratin fibrils to provide another form of crosslinkage.[i,j] Small proline-rich proteins, desmosonal proteins, and others are also present in the CE.[c]

As the final outer **stratum corneum** is formed the phospholipid bilayer deteriorates and intercellular lipid layers are formed.[k,l] These contain principally ceramides, cholesterol, and free fatty acids. Some sphingolipids are covalently attached to proteins.[a]

Also present in the epidermis are embedded macrophage-like Langerhans cells as well as pigmented **melanocytes**, cells with highly branched dendrites, which lie just above the basal stem cell layer. Each melanocyte contains hundreds of pigmented organelles called **melanosomes**. They contain not only the black or reddish **melanin** pigments but also the enzymes needed to form them (Chapter 25).[n,o]

The dendrites of a melanocyte contact about 36 keratinocytes and are able to transfer melanosomes to these adjacent cells. The numbers and sizes of the melanosomes as well as melanin structure determine differences in skin color.[m] Similar cells in amphibians, the **melanophores**, also contain light receptors.[p] Their melanosomes are not transferred to other cells but may be either clustered near the center of the cell or dispersed. The location can be changed quickly by transport of the melanosomes along a network of microtubules allowing the animals to change in response to changes in light color.[q] Various stimuli, including ultraviolet irradiation of melanocytes, cause increased synthesis of melanin with a resultant tanning[o] and added protection against sunburn.

Beneath the basement membrane of the epidermis is the **dermis**, a thick, tough, collagen-rich connective tissue. Blood vessels and nerve endings are found in this layer, as are roots of hairs and oil and sweat glands.[r]

Skin suffers from a variety of ailments including serious hereditary diseases.[a] One group of **keratinization disorders**, known as **ichthyoses**, are characterized by thickened, scaly skin. Some hereditary ichthyoses result from defects in type II

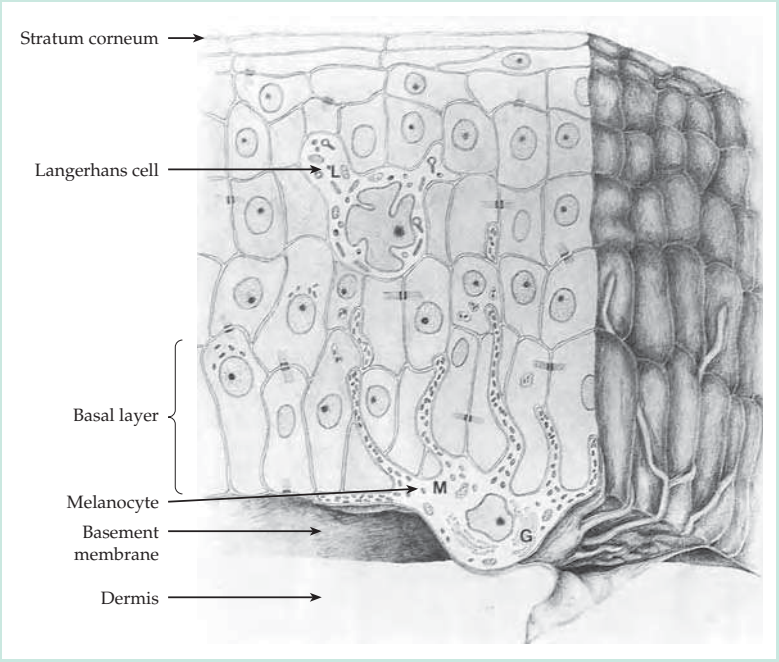



Diagram of a dendritic melanocyte surrounded by satellite keratinocytes. The Golgi area (G), where the melanosomes are synthesized, is shown around the nucleus. The other branched cell, higher in the epidermis, is a Langerhans cell with its tennis racquet-shaped granules. Courtesy of Dr. W. Quevedo, Jr. From Montagna *et al*.[m]

## BOX 8-F SKIN (continued)

keratin.[b] **Lamellar ichthyosis** reflects defects in the crosslinking enzyme transglutaminase.[e,f] **Epidermolysis bullosa** is a heterogeneous group of disorders characterized by easy formation of blisters. One form has been traced to a defect in the anchoring fibrils of type VII collagen, which tie cells of the basal layer to the basement membrane.[s,t] Others are defects in keratins of the basal or intermediate layers.[a] Yet others involve the lipid metabolism of skin, e.g. a steroid sulfatase deficiency.[a]

The most frequent skin disorder, which affects about 2% of the world's population is **psoriasis**. The thickened, scaly patches can cover much of the skin and become disabling. The inflammation and excessive epidermal growth are usually a T-cell mediated immunologic response to antigenic stimuli.[p,u,v] However, there is a hereditary form.[u] Other common skin disorders include actinic keratosis induced by light and cancer.


[a] Roop, D. (1995) *Science* **267**, 474–475
[b] Takahashi, K., Paladini, R. D., and Coulombe, P. A. (1995) *J. Biol. Chem.* **270**, 18581–18592
[c] Robinson, N. A., Lapic, S., Welter, J. F., and Eckert, R. L. (1997) *J. Biol. Chem.* **272**, 12035–12046
[d] Kim, S.-Y., Chung, S.-I., and Steinert, P. M. (1995) *J. Biol. Chem.* **270**, 18026–18035
[e] Huber, M., Rettler, I., Bernasconi, K., Frenk, E., Lavrijsen, S. P. M., Ponec, M., Bon, A., Lautenschlager, S., Schorderet, D. F., and Hohl, D. (1995) *Science* **267**, 525–528
[f] Candi, E., Melino, G., Lahm, A., Ceci, R., Rossi, A., Kim, I. G., Ciani, B., and Steinert, P. M. (1998) *J. Biol. Chem.* **273**, 13693–13702
[g] Yaffe, M. B., Beegen, H., and Eckert, R. L. (1992) *J. Biol. Chem.* **267**, 12233–12238
[h] Hohl, D., Mehrel, T., Lichti, U., Turner, M. L., Roop, D. R., and Steinert, P. M. (1991) *J. Biol. Chem.* **266**, 6626–6636
[i] Resing, K. A., Walsh, K. A., Haugen-Scofield, J., and Dale, B. A. (1989) *J. Biol. Chem.* **264**, 1837–1845
[j] Mack, J. W., Steven, A. C., and Steinert, P. M. (1993) *J. Mol. Biol.* **232**, 50–66
[k] ten Grotenhuis, E., Demel, R. A., Ponec, M., Boer, D. R., van Miltenburg, J. C., and Bouwstra, J. A. (1996) *Biophys. J.* **71**, 1389–1399
[l] Bouwstra, J. A., Thewalt, J., Gooris, G. S., and Kitson, N. (1997) *Biochemistry* **36**, 7717–7725
[m] Montagna, W., Prota, G., and Kenney, J. A., Jr. (1993) *Black Skin Structure and Function*, Academic Press, San Diego, California
[n] Potterf, S. B., Muller, J., Bernardini, I., Tietze, F., Kobayashi, T., Hearing, V. J., and Gahl, W. A. (1996) *J. Biol. Chem.* **271**, 4002–4008
[o] Roméro-Graillet, C., Aberdam, E., Biagoli, N., Massabni, W., Ortonne, J.-P., and Ballotti, R. (1996) *J. Biol. Chem.* **271**, 28052–28056
[p] Greaves, M. W., and Weinstein, G. D. (1995) *N. Engl. J. Med.* **332**, 581–588
[q] Rogers, S. L., Tint, I. S., Fanapour, P. C., and Gelfand, V. I. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3720–3725
[r] Martin, P. (1997) *Science* **276**, 75–81
[s] Byers, P. H. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4029–4077, McGraw-Hill, New York
[t] Christiano, A. M., Ryynänen, M., and Uitto, J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3549–3553
[u] Tomfohrde, J., Silverman, A., Barnes, R., Fernandez-Vina, M. A., Young, M., Lory, D., Morris, L., Wuepper, K. D., Stastny, P., Menter, A., and Bowcock, A. (1994) *Science* **264**, 1141–1145
[v] Boehncke, W.-H., Dressel, D., Zollner, T. M., and Kaufmann, R. (1996) *Nature (London)* **379**, 777


***Fibrillin and Marfan's syndrome.*** Most connective tissues contain insoluble, beaded microfibrils 10–12 nm in diameter. A component of some of these microfibrils, which are often found in elastic tissue, was purified from media used to culture human fibroblasts in 1986. This protein, called fibrillin, is a single-chain 350-kDa glycoprotein which contains ~14% cysteine.[673–674] Using a DNA probe based on the partially cloned fibrillin gene, the location of the gene was established on the long arm of chromosome 15 at a site previously identified as that of a gene defective in Marfan's syndrome. This disorder often causes dislocation of lenses of the eyes and aortic aneurysm as well as elongated limbs and fingers. Point mutations in the fibrillin gene have been identified in both Marfan's patients and family members that carry the defective gene.[675,676]

***The cuticles of invertebrates.*** The tough elastic cuticle of the nematode *Caenorhabditis* is largely collagen. However, the molecules are smaller than in vertebrates and there are ~100 different genes whose transcription gives rise to a large variety of similar proteins.[652] Cuticles of some annelids have unusually long collagens.[653,677] In contrast, the epithelial cells of insects and other arthropods secrete chitin which serves as the framework for development of a thick and often hard cuticle or exoskeleton. The cuticle also contains a variety of proteins.[678,679] In some instances mineralization by calcium carbonate occurs. During the later phases of the cuticle development extensive crosslinking of the proteins takes place. This is largely by reactions between modified aromatic side chains and resembles the chemistry of formation of melanin and lignin (Chapter 25).[680]

***Bones, teeth, and shells.*** Living organisms are able to induce the formation of over 60 inorganic compounds.[681] Most of these are formed by animals. Two forms of calcium carbonate, **calcite** and **aragonite**, predominate.[682] These minerals form shells, exoskeleton bones, bones, teeth, and other specialized structures.

While some organisms promote mineral deposition completely outside of their cell coats, the mineralization is usually controlled by the proteins and polysaccharides lying around and between cells.[683–687]

Bone deposition begins in the proteoglycan and collagen II matrix of cartilage. Later these polymers are largely replaced by collagen I which accounts for 90% of the organic material in mature bone. Bone collagen has a distinctive pattern of crosslinking.[639,688,689] Some borohydride-reducible crosslinks remain throughout adult life.[639] However, complex hydroxypyridine and pyrrole linkages are more characteristic of bone collagen. Embedded in the bone matrix are spidery cells, the **osteocytes**. Among these are **osteoblasts**,[689a] which secrete the collagen and other proteins that promote the laying down of calcium phosphate. Also present are large multinucleate **osteoclasts**[689b] which dissolve bone and reabsorb calcium and phosphate. Cells of both types remain active in mature bone, which is both a structural material and a store of calcium and phosphorus.

How do osteoblasts induce calcium phosphate deposition? The first crystals laid down within the cartilage matrix are of carbonate apatite only 2–3 nm thick and tens of nanometers in length. They are intimately associated with the collagen and other components of the matrix[690] and may be formed within matrix vesicles.[691] From observations with large biologically formed calcium carbonate crystals it is known that proteins or other organic initiators of crystallization can be found embedded in mature crystals. If the same is true in bone what are the initiators? The answer is uncertain but it is known that without osteoblasts the partially mineralized cartilage will not become bone.[692]

A key to the development of osteoblasts appears to be an **osteoblast-specific transcription factor OS/2** or **Cbfa1**.[693–695] Mutations in human Cbfa1 are linked to a series of skeletal defects.[696] A unique change accompanying conversion of a precursor cell into an osteoblast is the formation of a 49-residue γ-carboxyglutamate (Gla)-containing protein called **osteocalcin**.[697] (See also Box 15-F). It is the most abundant noncollagenous protein of bone. Its three Gla residues doubtless help to bind calcium ions and osteocalcin may be an initiator of crystallization. Osteocalcin has also been found in fish scales.[698] Also present in bone is a 74-residue **matrix Gla protein** which has 5 Gla residues.[699]

Other phosphoproteins, glycoproteins, and proteoglycans may also be required for mineralization. For example, the 286-residue glycoprotein **osteonectin**[323,700,701] accounts for 3% of total bone protein. It contains two $Ca^{2+}$-binding motifs and inhibits growth of hydroxylapatite crystals, but its role in bone development is not clear. One of the phosphoproteins of developing bone has been identified as

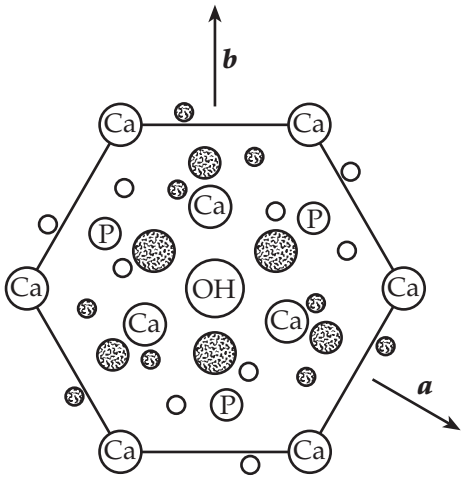



**Figure 8-34** An end-on view of a crystallite of hydroxylapatite. The shaded atoms of Ca, P, and O represent an underlying layer. The $OH^-$ groups form a longitudinal H-bonded array in the center. From J. A. Weatherell and C. Robinson,[705] p. 55. A small amount of $Mg^{2+}$ is present in place of $Ca^{2+}$ and a very small fraction of the $OH^-$ is replaced by fluoride ion which has a bone strengthening effect.

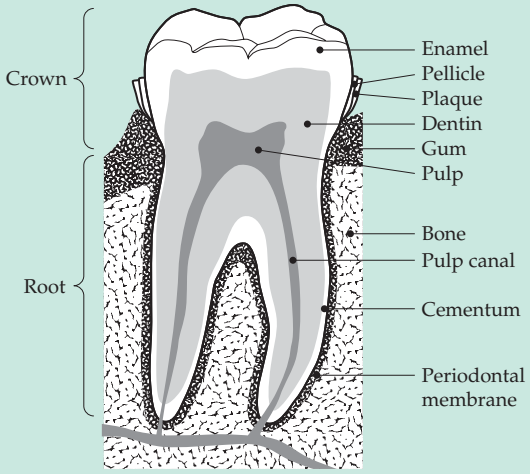
the 24 kDa propeptide cut from the N-terminal end of the α1 chain of type I procollagen.[702] Another bone protein, the sialic acid-containing **osteopontin** has the cell-binding sequence GRGDS identical to that in fibronectin.[703,704] Since it also binds to hydroxylapatite this protein may form a bridge between cell surfaces and the mineral matrix.

The mineral phase of bone is largely hydroxylapatite, $Ca_{10}(PO_4)_6(OH)_2$ (Fig. 8-34), which is essentially in chemical equilibrium with the calcium and phosphate ions present in the blood serum. Thus, bone cells can easily promote either the deposition or dissolution of the mineral phase by localized change in pH, concentrations of $Ca^{2+}$ or $HPO_4^{2-}$, or of chelating compounds such as ATP or inorganic pyrophosphate. Small 100 nm vesicles rich in acidic phospholipids and containing both $Ca^{2+}$ and the enzymes alkaline phosphatase and pyrophosphatase, may play an essential role in calcification. Perhaps they release the calcium and enzymes that generate inorganic phosphate.[691]

Bone is noted for its continual "remodeling" caused by the action of both osteoblasts and osteoclasts.[704a,b] The latter are large multinuclear cells derived from the same precursors as give rise to macrophages.[689b,704] Osteoclasts generate a sealed, acidic compartment on a bone surface, using a vacuolar ATPase to pump protons from the cytosol of the osteoclast, using $H_2CO_3$ as a source of protons.[706] The $HCO_3^-$ then leaves cells in exchange for $Cl^-$ which then also enters the acidified compartment. Not only is the calcium phosphate dissolved but lysomal enzymes digest the organic materials. Each osteoclast creates a long tun-

## BOX 8-G THE BIOCHEMISTRY OF TEETH

Like other bones, mammalian teeth are composed largely of collagen and hydroxylapatite. The much studied rat incisor is 65% mineral; about 85% of the organic material is type I collagen. A tooth consists of three mineralized tissues together with the internal blood vessels and nerves of the pulp.[a] The outer layer of **enamel** is formed by secretion from a thin layer of epithelial cells, the **ameloblasts**. The enamel matrix is devoid of collagen but contains two characteristic groups of proteins, the **amelogenins** and the **enamelins**.




This cross section of a human molar shows deposits of pellicle and plaque near the gum line[h]

The amelogenins are hydrophobic but are also rich in proline, histidine, and glutamic acid.[a,b] They account for 90% of the matrix proteins but are replaced by the initially less abundant enamelins[c] in fully mineralized teeth.[a,d] Another protein present in developing enamel is **ameloblastin**, which appears to be unique to amyloblasts.[a] As mineralization of the enamel progresses the matrix is lost and 98% of the enamel is hydroxylapatite.[a]

Dentin is of epithelial-mesenchymal origin. The **odontoblasts** secrete an extracellular matrix that is rich in collagen I and which also contains all of the other major bone proteins as well as a **dentin sialoprotein**[e] and poorly characterized phosphate carriers, the **phosphophoryns**.

The third hard tissue of teeth is **cementum**, which binds teeth to the periodental ligaments. The ligaments contain fibrillar collagen which inserts into the cementum.[f,g]

Except for the common cold, tooth decay (**caries**) is the most prevalent disease in the United States.[h,i] Caries is initiated by attack of acids produced by bacterial fermentations on the enamel. The saliva contains calcium and phosphate and is supersaturated with respect to these ions. As a result the enamel surface is continuously recalcified. The 43-residue salivary protein **statherin** retards precipitation of calcium phosphates from the saliva, preventing excessive build-up of calcified deposits on the teeth.[j] Other proline-rich proteins may play a role in recalcification repair.[k]

A freshly cleaned tooth surface quickly becomes coated with a thin **pellicle** of salivary proteins. This provides a surface for growth of **dental plaque**, which contains many bacteria and adhesive polysaccharides such as dextrans.[l] The latter are generated from dietary sucrose by such bacteria as *Streptococcus mutans*. (Chapter 20) and others.[m] Many factors affect the probability of tooth decay. A high sucrose diet promotes decay.[n] While most people have some trouble with tooth decay, 1 or 2 per thousand remain totally free of caries and seem to be immune. Many factors must affect resistance to caries. For example, individuals vary in the kinds and numbers of bacteria present on teeth and in the structure of tooth enamel.[o] Addition of fluoride ion to water supplies at a level of 1ppm (0.05 mM) is generally believed to reduce the incidence of tooth decay. However, caries has been declining in many developed countries at rates that are the same for water with or without fluoride.[p,q] If teeth escape caries **periodental disease**, caused by bacteria, is often a major problem for older people.[r]


[a] Krebsbach, P. H., Lee, S. K., Matsuki, Y., Kozak, C. A., Yamada, K. M., and Yamada, Y. (1996) *J. Biol. Chem.* **271**, 4431–4435
[b] Renugopalakrishnan, V., Strawich, E. S., Horowitz, P. M., and Glimcher, M. J. (1986) *Biochemistry* **25**, 4879–4887
[c] Deutsch, D., Palmon, A., Fisher, L. W., Kolodny, N., Termine, J. D., and Young, M. F. (1991) *J. Biol. Chem.* **266**, 16021–16028
[d] Robinson, C., Weatherell, J. A., and Höhling, H. J. (1983) *Trends Biochem. Sci.* **8**, 284–287
[e] Ritchie, H. H., Hou, H., Veis, A., and Butler, W. T. (1994) *J. Biol. Chem.* **269**, 3698–3702
[f] Arzate, H., Olson, S. W., Page, R. C., Gown, A. M., and Narayanan, A. S. (1992) *FASEB J.* **6**, 2990–2995
[g] Yamauchi, M., Katz, E. P., and Mechanic, G. L. (1986) *Biochemistry* **25**, 4907–4913
[h] Sanders, H. J. (1980) *Chem. Eng. News* **(Feb. 25) 58(8)** 30–39
[i] Shaw, J. H. (1987) *N. Engl. J. Med.* **317**, 996–1004
[j] Schlesinger, D. H., and Hoy, D. I. (1977) *J. Biol. Chem.* **252**, 1689–1695
[k] Clements, S., Mehansho, H., and Carlson, D. M. (1985) *J. Biol. Chem.* **260**, 13471–13477
[l] Kolenbrander, P. E., Ganeshkumar, N., Cassels, F. J., and Hughes, C. V. (1993) *FASEB J.* **7**, 406–408
[m] Abeygunawardana, C., and Bush, C. A. (1990) *Biochemistry* **29**, 234–248
[n] Newbrun, E. (1982) *Science* **217**, 418–423
[o] Cevc, G., Cevc, P., Schara, M., and Skaleric, U. (1980) *Nature (London)* **286**, 425–426
[p] Diesendorf, M. (1986) *Nature (London)* **322**, 125–129
[q] Hileman, B. (1988) *Chem. Eng. News* **Aug. 1**, 26–42
[r] Williams, R. C. (1990) *N. Engl. J. Med.* **322**, 373–382

nel into the interior of the bone. However, the tunnel is soon lined with new osteoblasts and new bone is laid down.[707] In **osteoporosis**, a very common disease of older persons, the rate of resorption of bone by osteoclasts exceeds that of bone formation by osteoblasts. Women typically lose 50% and men ~30% of their calcium phosphate from vertebrae and ends of long bones as they age.[708] In osteopetrosis (marble bone disease) excessive calcification of bone and other tissues occurs as a result of a deficiency of carbonic anhydrase.[709] See also Chapter 13. A disease of progressive calcification of soft tissues[710] results from an excess of **bone morphogenic factor-4**, one of a group of protein factors acting on bone development (Chapter 32).[631] **Paget's disease** is another disorder of bone remodeling that leads to overproduction of bone of poor quality.[711] As discussed in Box 22-C, the circulating level of $Ca^{2+}$ as well as the cellular uptake of this ion are controlled by vitamin D and its metabolites and by calcitonin and the parathyroid hormones.

Lack of a pyrophosphate ion pump in cartilage cells may cause a deficit in pyrophosphate in the surroundings of the forming bone. In mice with a defect in this pump bony spurs, similar to those in human osteoarthritis, are formed.[711a] Over half of the world's population of persons over 65 years of age are afflicted by arthritis, over 100 types being known.[711b]

Calcium carbonate in several different crystalline forms arises biologically.[705,712] Sometimes the mineral is deposited within vesicles inside the cell. Thus some species of the protozoan *Hymenomonas* use Golgi vesicles as sites for construction of segments of plantlike cell walls complete with crystalline calcium carbonate plates attached to the outer surface.[713] These wall segments are then transported to their final locations. The intricately sculptured spicules of sea urchins and sponges are single crystals of calcium carbonate which have grown within intracellular vesicles.[687,714] Diatoms also form their shells of nearly pure, hydrated $SiO_2$ (See Box 4-B) entirely within membrane-bound vesicles.[715] On the other hand, shells of molluscs are usually deposited outside the cells of the organism but again under the influence of a protein matrix.[716,717] The animals apparently actively pump bicarbonate outward where it reacts with $Ca^{2+}$ (Eq. 8-11).[718]

$$Ca^{2+} + HCO_3^- \rightarrow CaCO_3 + H^+ \qquad (8\text{-}11)$$

The protons released then react with more bicarbonate to form carbonic acid which is converted to $CO_2$ and water (Eq. 8-12) by the enzyme carbonic anhydrase.

$$H_2CO_3 \xrightarrow{\text{Carbonic anhydrase}} CO_2 + H_2O \qquad (8\text{-}12)$$

A problem that arises within organisms is avoiding crystallization under supersaturating conditions. For example, normal urine is supersaturated with calcium oxalate. To prevent formation of renal calculi (stones)[719] an inhibitory glycoprotein is present and slows the formation and growth of crystals.[720] Under some disease conditions calcium carbonate stones may form in pancreatic ducts. A 17 kDa lectinlike glycoprotein called **lithostatine** has been proposed to inhibit stone formation by binding to certain planes on $CaCO_3$ microcrystals just as antifreeze proteins (Box 4-D) inhibit ice formation.[721] However, this proposed function for lithostatine is doubtful.[722,723] Pathological deposits of crystalline calcium pyrophosphate and basic calcium phosphates are sometimes present in joints,[724] even in Neanderthal skeletons.[725]

## 3. Cell Walls of Fungi and Green Plants

The cell walls of yeasts and other fungi are made up largely of glucans, chitin, and a mannan-protein complex. Yeasts contain predominately glucans but chitin is the major polysaccharide in many other fungi. The most abundant glucan is a β1,3 linked polymer with about 3% β1,6 linkages to the branches and a molecular mass of ~ 240 kDa. In addition there is about 15% of a highly branched 1,6 linked glucan containing 1,3 linked branches. At the outer surface of the wall are mannoproteins[726–728] which carry small serine- or threonine-linked oligomannans as well as large highly branched mannans linked to asparagine through the usual *N*-acetylglucosamine-containing core structure. Some of these highly branched mannan chains serve as species-specific antigens.[727] Like those of the bacterial and animal cell surfaces, the antigens vary in structure, a fact with important medical implications. Curiously, many fungi have surface fimbriae which are composed of **collagen**, which is usually regarded as exclusively an animal protein.[729]

Cell walls of higher plants (Fig. 1-7, 4-14) are composed largely of polysaccharides. They are discussed briefly in Chapter 20.

## References


1. Christie, W. W. (1973) *Lipid Analysis; Isolation, Preparation, Identification and Structural Analysis of Lipids*, Pergamon, Oxford
2. Hitchcock, C., and Nichols, B. W. (1971) *Plant Lipid Biochemistry*, Academic Press, New York
3. Kates, M. (1972) in *Laboratory Techniques in Biochemistry and Molecular Biology*, Vol. 3, Part II (Work, T. S., and Work, E., eds), North-Holland Publ., Amsterdam
4. Weete, J. W. (1980) *Lipid Biochemistry of Fungi and Other Organisms*, Plenum, New York
5. Gurr, M. I., and James, A. T. (1975) *Lipid Biochemistry, an Introduction*, 2nd ed., Cornell Univ. Press, Ithaca, New York
6. Porter, K. R., and Bonneville, M. A. (1973) *Fine Structure of Cells and Tissues*, 4th ed., Lea & Febiger, Philadelphia, Pennsylvania
7. Aveldaño, M. I., and Sprecher, H. (1987) *J. Biol. Chem.* **262**, 1180–1186
8. Gunstone, F. D. (1967) *An Introduction to the Chemistry and Biochemistry of Fatty Acids and Their Glycerides*, 2nd ed., Chapman and Hall, London (Chapter 6)
9. Steck, P. A., Schwartz, B. A., Rosendahl, M. S., and Gray, G. R. (1978) *J. Biol. Chem.* **253**, 5625–5629
10. McNeil, M., Daffe, M., and Brennan, P. J. (1991) *J. Biol. Chem.* **266**, 13217–13223
11. Qureshi, N., Takayama, K., and Schnoes, H. K. (1980) *J. Biol. Chem.* **255**, 182–189

11a. Adlof, R. O., Duval, S., and Emken, E. A. (2000) *Lipids* **35**, 131–135

12. Cravatt, B. F., Prospero-Garcia, O., Siuzdak, G., Gilula, N. B., Henriksen, S. J., Boger, D. L., and Lerner, R. A. (1995) *Science* **268**, 1506–1509
13. Hummel, H. E., Gaston, L. K., Shorey, H. H., Kaae, R. S., Byrne, K. J., and Silverstein, R. M. (1973) *Science* **181**, 873–874
14. Klun, J. A., Chapman, O. L., Mattes, K. C., Wojtkowski, P. W., Beroza, M., and Sonnet, P. E. (1973) *Science* **181**, 661–663
15. Moolenaar, W. H. (1995) *J. Biol. Chem.* **270**, 12949–12952
16. Wells, W. W., and Eisenberg, F., Jr., eds. (1978) *Cyclitols and Phosphoinositides*, Academic Press, New York
17. DeBuch, H. (1978) *Trends Biochem. Sci.* **3**, 44–45
18. Chilton, F. H., and Connell, T. R. (1988) *J. Biol. Chem.* **263**, 5260–5265
19. Ford, D. A., and Gross, R. W. (1994) *Biochemistry* **33**, 1216–1222
20. Friedman, S. M., ed. (1978) *Biochemistry of Thermophily*, Academic Press, New York
21. Kushwaha, S. C., Kates, M., Sprott, G. D., and Smith, I. C. P. (1981) *Science* **211**, 1163–1164
22. Luzzati, V., Gambacorta, A., DeRosa, M., and Gulik, A. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 25–47
23. Goldfine, H., and Langworthy, T. A. (1988) *Trends Biochem. Sci.* **13**, 217–221
24. Ferrante, G., Ekiel, I., and Sprott, G. D. (1986) *J. Biol. Chem.* **261**, 17062–17066
25. Pond, J. L., Langworthy, T. A., and Holzer, G. (1986) *Science* **231**, 1134–1136
26. Florin-Christensen, J., Florin-Christensen, M., Knudson, J., and Rasmussen, L. (1986) *Trends Biochem. Sci.* **11**, 354–355
27. Cooney, R. V., Mumma, R. O., and Benson, A. A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4262–4264
28. Renou, J.-P., Giziewicz, J. B., Smith, I. C. P., and Jarrell, H. C. (1989) *Biochemistry* **28**, 1804–1814
29. Gounaris, K., and Barber, J. (1983) *Trends Biochem. Sci.* **8**, 378–381
30. Knowles, F. C., and Benson, A. A. (1983) *Trends Biochem. Sci.* **8**, 178–180
31. Tupper, S., Wong, P. T. T., Kates, M., and Tanphaichitr, N. (1994) *Biochemistry* **33**, 13250–13258
32. Andersson, A.-S., Rilfors, L., Bergqvist, M., Persson, S., and Lindblom, G. (1996) *Biochemistry* **35**, 11119–11130

32a. Vikström, S., Li, L., and Wieslander, Å. (2000) *J. Biol. Chem.* **275**, 9296–9302

33. Nishihara, M., Utagawa, M., Akutsu, H., and Koga, Y. (1992) *J. Biol. Chem.* **267**, 12432–12435
34. Powles, T. J., Easty, D. M., Easty, G. C., Bondy, P. K., and Munro-Neville, A. (1973) *Nature (London), New Biol.* **245**, 83
35. Lynch, D. V. (1993) in *Lipid Metabolism in Plants* (Moore, T. S., Jr., ed), pp. 285–308, CRC Press, Boca Raton, Florida
36. Araki, S., Satake, M., Ando, S., Hayashi, A., and Fujii, N. (1986) *J. Biol. Chem.* **261**, 5138–5144
37. Araki, S., Abe, S., Odani, S., Ando, S., Fujii, N., and Satake, M. (1987) *J. Biol. Chem.* **262**, 14141–14145
38. Song, W., and Rintoul, D. A. (1989) *Biochemistry* **28**, 4194–4200
39. Hsieh, T. C.-Y., Lester, R. L., and Laine, R. A. (1981) *J. Biol. Chem.* **256**, 7747–7755
40. Barr, K., and Lester, R. L. (1984) *Biochemistry* **23**, 5581–5588
41. Dickson, R. C., Nagiec, E. E., Wells, G. B., Nagiec, M. M., and Lester, R. L. (1997) *J. Biol. Chem.* **272**, 29620–29625
42. Yamakawa, T., and Nagai, Y. (1978) *Trends Biochem. Sci.* **3**, 128–131
43. Hansson, G. C. (1983) *J. Biol. Chem.* **258**, 9612–9615
44. Clausen, H., Holmes, E., and Hakomori, S.-I. (1986) *J. Biol. Chem.* **261**, 1388–1392
45. Farooqui, A. A. (1981) *Adv. Lipid Res.* **18**, 159–202
46. Godchaux, W., III, and Leadbetter, E. R. (1984) *J. Biol. Chem.* **259**, 2982–2990
47. Abbanat, D. R., Leadbetter, E. R., Godchaux, W., III, and Escher, A. (1986) *Nature (London)* **324**, 367–369
48. Jones, M. N., and Chapman, D., eds. (1994) *Micelles, Monolayers and Biomembranes*, Wiley, New York
49. Jain, M. (1988) *Introduction to Biological Membranes*, 2nd ed., Wiley, New York
50. Lofan, R., and Nicolson, G. L. (1981) in *Advanced Cell Biology* (Schwartz, L. M., and Azar, M. M., eds), Van Nostrand Reinhold Co., New York
51. Storch, J., and Kleinfeld, A. M. (1985) *Trends Biochem. Sci.* **10**, 418–421
52. Robertson, R. N. (1983) *The Lively Membranes*, Cambridge Univ. Press, Cambridge
53. Starzak, M. (1984) *The Physical Chemistry of Membranes*, Academic Press, San Diego, California
54. Gennis, R. B. (1989) *Biomembranes: Molecular Structure and Function*, Springer-Verlag, New York
55. Morell, P., and Norton, W. T. (1980) *Sci. Am.* **242**(May), 88–116
56. Martenson, R. E., ed. (1992) *Myelin: Biology and Chemistry*, CRC Press, Boca Raton, Florida
57. Weimbs, T., and Stoffel, W. (1994) *Biochemistry* **33**, 10408–10415
58. Korn, E. D. (1966) *Science* **153**, 1491–1498
59. Tanford, C. (1980) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, 2nd ed., Wiley, New York
60. Danielli, J. F., and Davsen, H. (1985) *J. Cell. Comp. Physiol.* **5**, 495–508
61. Singer, S. J., and Nicolson, G. L. (1972) *Science* **175**, 720–731
62. Bretscher, M. S., and Raff, M. C. (1975) *Nature (London)* **258**, 43–49
63. Quinn, P. J., and Cherry, R. J., eds. (1992) *Structural and Dynamic Properties of Lipids and Membranes*, Portland Press, London
64. Disalvo, E. A., and Simon, S. A., eds. (1995) *Permeability and Stability of Lipid Bilayers*, CRC Press, Boca Raton, Florida
65. Jacobson, K., Sheets, E. D., and Simson, R. (1995) *Science* **268**, 1441–1442
66. Pascher, I., Lundmark, M., Nyholm, P.-G., and Sundell, S. (1992) *Biochim. Biophys. Acta.* **1113**, 339–373
67. Cameron, D. G., Gudgin, E. F., and Mantsch, H. H. (1981) *Biochemistry* **20**, 4496–4500
68. Zasadzinski, J. A. N., Helm, C. A., Longo, M. L., Weisenhorn, A. L., Gould, S. A. C., and Hansma, P. K. (1991) *Biophys. J.* **59**, 755–760
69. Wiener, M. C., and White, S. H. (1992) *Biophys. J.* **61**, 428–433
70. Smith, D. P. E., Bryant, A., Quate, C. F., Rabe, J. P., Gerber, C. H., and Swalen, J. D. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 969–972
71. Yang, J., Tamm, L. K., Tillack, T. W., and Shao, Z. (1993) *J. Mol. Biol.* **229**, 286–290
72. Nagle, J. F., Zhang, R., Tristram-Nagle, S., Sun, W., Petrache, H. I., and Suter, R. M. (1996) *Biophys. J.* **70**, 1419–1431
73. Mueller, P., Rudin, D. O., Tien, H. T., and Wescott, W. C. (1962) *Nature (London)* **194**, 979–980
74. Howard, K. P., and Prestegard, J. H. (1995) *J. Am. Chem. Soc.* **117**, 5031–5040
75. Ostro, M. J. (1987) *Sci. Am.* **256**(Jan), 103–111
76. Philippot, J. R., and Schuber, F., eds. (1995) *Liposomes as Tools in Basic Research and Industry*, CRC Press, Boca Raton, Florida
77. Harwood, J. L. (1992) *Trends Biochem. Sci.* **17**, 203–204

77a. Schröder, M. (1980) *Chem. Rev.* **80**, 187–213

77b. Mohapatra, S. K., and Behrman, E. J. (1982) *J. Inorg. Biochem* **16**, 85–89

77c. Deetz, J. S., and Behrman, E. J. (1981) *Intl. J. Peptide Prot. Res.* **17**, 495–500

78. Norrby, P.-O., Becker, H., and Sharpless, K. B. (1996) *J. Am. Chem. Soc.* **118**, 35–42
79. Montal, M., and Mueller, P. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3561–3566
80. Ladha, S., Mackie, A. R., Harvey, L. J., Clark, D. C., Lea, E. J. A., Brullemans, M., and Duclohier, H. (1996) *Biophys. J.* **71**, 1364–1373
81. Jones, M. N., and Chapman, D. (1995) *Micelles, Monolayers, and Biomembranes*, Wiley-Liss, New York

81a. Knobler, C. M. (1990) *Adv. Chem. Phys.* **77**, 397–449

82. Bryce, M. R., and Petty, M. C. (1995) *Nature (London)* **374**, 771–776
83. Zasadzinski, J. A., Viswanathan, R., Madsen, L., Garnaes, J., and Schwartz, D. K. (1994) *Science* **263**, 1726–1733
84. Fettiplace, R., Andrews, D. M., and Haydon, D. A. (1971) *J. Membr. Biol.* **5**, 277–296
85. White, S. W., and King, G. I. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6532–6536
86. Snyder, R. G., Maroncelli, M., Qi, S. P., and Strauss, H. L. (1981) *Science* **214**, 188–190
87. Di, L., and Small, D. M. (1995) *Biochemistry* **34**, 16672–16677
88. Tu, K., Tobias, D. J., Blasie, J. K., and Klein, M. L. (1996) *Biophys. J.* **70**, 595–608
89. Katsaras, J., Raghunathan, V. A., Dufourc, E. J., and Dufourcq, J. (1995) *Biochemistry* **34**, 4684–4688
90. Snyder, R. G., Liang, G. L., Strauss, H. L., and Mendelsohn, R. (1996) *Biophys. J.* **71**, 3186–3198
91. Chen, S. C., Sturtevant, J. M., and Gaffney, B. J. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5060–5063

## References


92. Sun, W.-J., Tristram-Nagle, S., Suter, R. M., and Nagle, J. F. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7008–7012
93. Mou, J., Yang, J., and Shao, Z. (1994) *Biochemistry* **33**, 4439–4443
94. Wang, Z.-q, Lin, H.-n, Li, S., and Huang, C.-h. (1995) *J. Biol. Chem.* **270**, 2014–2023
95. Wang, G., Lin, H.-n, Li, S., and Huang, C.-h. (1995) *J. Biol. Chem.* **270**, 22738–22746
96. Wang, Z.-q, Lin, H.-n, Li, S., and Huang, C.-h. (1994) *J. Biol. Chem.* **269**, 23491–23499
97. Dufourc, E. J., Parish, E. J., Chitrakorn, S., and Smith, I. C. P. (1984) *Biochemistry* **23**, 6062–6071
98. Maulik, P. R., and Shipley, G. G. (1996) *Biochemistry* **35**, 8025–8034
99. McMullen, T. P. W., and McElhaney, R. N. (1997) *Biochemistry* **36**, 4979–4986
100. Finegold, L., ed. (1993) *Cholesterol in Membrane Models*, CRC Press, Boca Raton, Florida
101. Recktenwald, D. J., and McConnell, H. M. (1981) *Biochemistry* **20**, 4505–4510
102. Needham, D., McIntosh, T. J., and Evans, E. E. (1988) *Biochemistry* **27**, 4668–4673
103. Goldfine, H., Johnston, N. C., and Phillips, M. C. (1981) *Biochemistry* **20**, 2908–2916
104. Lewis, R. V. A. H., Pohle, W., and McElhaney, R. N. (1996) *Biophys. J.* **70**, 2736–2746
105. Shimshick, E. J., and McConnell, H. M. (1973) *Biochemistry* **12**, 2351–2360
106. Almeida, P. F. F., Vaz, W. L. C., and Thompson, T. E. (1992) *Biochemistry* **31**, 7198–7210
107. Träuble, H., and Eibl, H. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 214–219
108. Basánez, G., Nieva, J. L., Rivas, E., Alonso, A., and Goni, F. M. (1996) *Biophys. J.* **70**, 2299–2306
109. Smith, R. L., and Oldfield, E. (1984) *Science* **225**, 280–288
109a. Smith, I. C. P. (1983) in *NMR of Newly Accessible Nuclei*, Vol. 2 (Laslo, P., ed), pp. 1–26, Academic Press, New York
110. Marassi, F. M., and Macdonald, P. M. (1992) *Biochemistry* **31**, 10031–10036
111. Oldfield, E., Bowers, J. L., and Forbes, J. (1987) *Biochemistry* **26**, 6919–6923
112. Bruzik, K. S., and Nyholm, P.-G. (1997) *Biochemistry* **36**, 566–575
113. Hong, M., Schmidt-Rohr, K., and Nanz, D. (1995) *Biophys. J.* **69**, 1939–1950
114. Thurmond, R. L., Niemi, A. R., Lindblom, G., Wieslander, Ä., and Rilfors, L. (1994) *Biochemistry* **33**, 13178–13188
115. Boden, N., Jones, S. A., and Sixl, F. (1991) *Biochemistry* **30**, 2146–2155
116. Davis, J. H. (1993) in *Cholesterol in Membrane Models* (Finegold, L., ed), pp. 67–157, CRC Press, Boca Raton, Florida
117. Baenziger, J. E., Jarrell, H. C., and Smith, I. C. P. (1992) *Biochemistry* **31**, 3377–3385
118. De Boeck, H., and Zidovetzki, R. (1992) *Biochemistry* **31**, 623–630
119. de Kruijft, B., Cullis, P. R., and Verkleij, A. J. (1980) *Trends Biochem. Sci.* **5**, 79–81
120. Peng, X., and Jonas, J. (1992) *Biochemistry* **31**, 6383–6390
121. Hong, M., Schmidt-Rohr, K., and Zimmermann, H. (1996) *Biochemistry* **35**, 8335–8341
122. Lu, D., Singh, D., Morrow, M. R., and Grant, C. W. M. (1993) *Biochemistry* **32**, 290–297
123. Salmon, A., and Hamilton, J. A. (1995) *Biochemistry* **34**, 16065–16073
124. Ruocco, M. J., Siminovitch, D. J., Long, J. R., Das Gupta, S. K., and Griffin, R. G. (1996) *Biophys. J.* **71**, 1776–1788
125. Bruzik, K. S., and Harwood, J. S. (1997) *J. Am. Chem. Soc.* **119**, 6629–6637
126. Casal, H. L., and McElhaney, R. N. (1990) *Biochemistry* **29**, 5423–5427
127. Hübner, W., Mantsch, H. H., Paltauf, F., and Hauser, H. (1994) *Biochemistry* **33**, 320–326
128. Léger, C. L., Daveloose, D., Christon, R., and Viret, J. (1990) *Biochemistry* **29**, 7269–7275
129. Jakobsson, E. (1997) *Trends Biochem. Sci.* **22**, 339–344
130. Berger, O., Edholm, O., and Jähnig, F. (1997) *Biophys. J.* **72**, 2002–2013
131. Chiu, S.-W., Clark, M., Balaji, V., Subramaniam, S., Scott, H. L., and Jakobsson, E. (1995) *Biophys. J.* **69**, 1230–1245
132. Browning, J. L. (1981) *Biochemistry* **20**, 7144–7151
133. Cevc, G., Watts, A., and Marsh, D. (1981) *Biochemistry* **20**, 4955–4965
134. Harlos, K., and Eibl, H. (1980) *Biochemistry* **19**, 895–899
135. Koynova, R. D., Kuttenreich, H. L., Tenchov, B. G., and Hinz, H.-J. (1988) *Biochemistry* **27**, 4612–4619
136. Hayakawa, E., Naganuma, M., Mukasa, K., Shimozawa, T., and Araiso, T. (1998) *Biophys. J.* **74**, 892–898
137. Katsaras, J., Jeffrey, K. R., Yang, D. S.-C., and Epand, R. M. (1993) *Biochemistry* **32**, 10700–10707
138. Srisiri, W., Sisson, T. M., O'Brien, D. F., McGrath, K. M., Han, Y., and Gruner, S. M. (1997) *J. Am. Chem. Soc.* **119**, 4866–4873
139. Siegel, D. P., and Epand, R. M. (1997) *Biophys. J.* **73**, 3089–3111
140. Thurmond, R. L., Lindblom, G., and Brown, M. F. (1993) *Biochemistry* **32**, 5394–5410
141. Siegel, D. P., Banschbach, J., and Yeagle, P. L. (1989) *Biochemistry* **28**, 5010–5019
142. Siegel, D. P., and Banschbach, J. L. (1990) *Biochemistry* **29**, 5975–5981
143. Landau, E. M., and Rosenbusch, J. P. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14532–14535
144. Delacroix, H., Gulik-Krzywicki, T., and Seddon, J. M. (1996) *J. Mol. Biol.* **258**, 88–103
144a. Tenchov, B., Koynova, R., and Rapp, G. (1998) *Biophys. J.* **75**, 853–866
145. Russell, N. J. (1984) *Trends Biochem. Sci.* **9**, 108–112
146. deMendoza, D., and Cronan, J. E., Jr. (1983) *Trends Biochem. Sci.* **8**, 49–52
147. Mantsch, H. H., Madec, C., Lewis, R. N. A. H., and McElhaney, R. N. (1987) *Biochemistry* **26**, 4045–4049
148. Bechtel, D. B., Mueller, D. D., Whaley, T. W., and Bulla, L. A., Jr. (1985) *J. Biol. Chem.* **260**, 9784–9792
149. Bérubé, L. R., and Hollingsworth, R. I. (1995) *Biochemistry* **34**, 12005–12011
150. Huster, D., Jin, A. J., Arnold, K., and Gawrisch, K. (1997) *Biophys. J.* **73**, 855–864
151. Träuble, H. (1971) *J. Membr. Biol.* **4**, 193–208
152. Wilson, M. A., and Pohorille, A. (1996) *J. Am. Chem. Soc.* **118**, 6580–6587
153. Ho, C., and Stubbs, C. D. (1997) *Biochemistry* **36**, 10630–10637
154. Bretscher, M. S. (1980) *Trends Biochem. Sci.* **5**, VI–VII
155. Devaux, P., and McConnell, H. M. (1972) *J. Am. Chem. Soc.* **94**, 4475–4481
156. Kell, D. B. (1984) *Trends Biochem. Sci.* **9**, 86–88
157. Picard, F., Paquet, M.-J., Dufourc, E. J., and Auger, M. (1998) *Biophys. J.* **74**, 857–868
158. Schmidt, T., Schütz, G. J., Baumgartner, W., Gruber, H. J., and Schindler, H. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2926–2929
159. Sheets, E. D., Lee, G. M., Simson, R., and Jacobson, K. (1997) *Biochemistry* **36**, 12449–12458
160. Hopkins, C. R. (1992) *Trends Biochem. Sci.* **17**, 27–32
161. Kleinfeld, A. M., Chu, P., and Storch, J. (1997) *Biochemistry* **36**, 5702–5711
162. Zhao, J., Zhou, Q., Wiedmer, T., and Sims, P. J. (1998) *Biochemistry* **37**, 6361–6366
163. Wolff, D., Canessa-Fischer, M., Vargas, F., and Diaz, G. (1971) *J. Membr. Biol.* **6**, 304–314
164. Sherbet, G. V. (1978) *The Biophysical Characterisation of the Cell Surface*, Academic Press, London
165. Jones, M. N. (1975) *Biological Interfaces*, Elsevier, Amsterdam
166. McLaughlin, S. (1989) *Ann. Rev. Biophys. Biophys. Chem.* **18**, 113–136
167. Aidley, D. J. (1971) *The Physiology of Excitable Cells*, Cambridge Univ. Press, London and New York
168. Nystrom, R. A. (1973) *Membrane Physiology*, Prentice-Hall, Englewood Cliffs, New Jersey
169. Nicholls, D. G., and Ferguson, S. J. (1992) *Bioenergetics 2*, Academic Press, London
170. Wood, P. M. (1985) *Trends Biochem. Sci.* **10**, 106–107
171. Teissié, J., Prats, M., LeMassu, A., Stewart, L. C., and Kates, M. (1990) *Biochemistry* **29**, 59–65
172. Teissié, J., Gabriel, B., and Prats, M. (1993) *Trends Biochem. Sci.* **18**, 243–246
173. Heberle, J., Riesle, J., Thiedemann, G., Oesterhelt, D., and Dencher, N. A. (1994) *Nature (London)* **370**, 379– 382
174. Gabriel, B., and Teissié, J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14521–14525
175. Devaux, P. F. (1991) *Biochemistry* **30**, 1163–1173
176. Victorov, A. V., Janes, N., Taraschi, T. F., and Hoek, J. B. (1997) *Biophys. J.* **72**, 2588–2598
177. Shiffer, K. A., Rood, L., Emerson, R. K., and Kuypers, F. A. (1998) *Biochemistry* **37**, 3449–3458
178. Michel, H. (1983) *Trends Biochem. Sci.* **8**, 56–59
179. Cowan, S. W., and Rosenbusch, J. P. (1994) *Science* **264**, 914–916
180. von Heijne, G. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 167–192
181. Bloom, M. (1995) *Biophys. J.* **69**, 1631–1632
182. Weimbs, T., and Stoffel, W. (1992) *Biochemistry* **31**, 12289–12296
183. Yamaguchi, Y., Ikenaka, K., Niinobe, M., Yamada, H., and Mikoshiba, K. (1996) *J. Biol. Chem.* **271**, 27838–27846
184. Jennings, M. L. (1989) *Ann. Rev. Biochem.* **58**, 999–1027
185. Bowie, J. U. (1997) *J. Mol. Biol.* **272**, 780–789
186. Gafvelin, G., Sakaguchi, M., Andersson, H., and von Heijne, G. (1997) *J. Biol. Chem.* **272**, 6119–6127
187. Jähnig, F. (1990) *Trends Biochem. Sci.* **15**, 93–95
188. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1994) *Biochemistry* **33**, 3038–3049
189. Persson, B., and Argos, P. (1996) *Protein Sci.* **5**, 363–371
190. Fasman, G. D., and Gilbert, W. A. (1990) *Trends Biochem. Sci.* **15**, 89–92
191. Milpetz, F., Argos, P., and Persson, B. (1995) *Trends Biochem. Sci.* **20**, 204–205
192. Shai, Y. (1995) *Trends Biochem. Sci.* **20**, 460–464
192a. Killian, J. A., and von Heijne, G. (2000) *Trends Biochem. Sci.* **25**, 429–434
192b. Lew, S., Ren, J., and London, E. (2000) *Biochemistry* **39**, 9632–9640
193. Popov, M., Tam, L. Y., Li, J., and Reithmeier, R. A. F. (1997) *J. Biol. Chem.* **272**, 18325–18332
194. Lee, G. F., Burrows, G. G., Lebert, M. R., Dutton, D. P., and Hazelbauer, G. L. (1994) *J. Biol. Chem.* **269**, 29920–29927
195. Lee, G. F., Dutton, D. P., and Hazelbauer, G. L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5416–5420
196. Yu, H., Kono, M., McKee, T. D., and Oprian, D. D. (1995) *Biochemistry* **34**, 14963–14969
197. Sansom, M. S. P., Son, H. S., Sankararamakrishnan, R., Kerr, I. D., and Breed, J. (1995) *Biophys. J.* **68**, 1295–1310

## References


198. Stopar, D., Jansen, K. A. J., Páli, T., Marsh, D., and Hemminga, M. A. (1997) *Biochemistry* **36**, 8261–8268
199. Nikaido, H. (1994) *J. Biol. Chem.* **269**, 3905–3908
200. Mishra, V. K., and Palgunachari, M. N. (1996) *Biochemistry* **35**, 11210–11220
201. Mishra, V. K., Palgunachari, M. N., Lund-Katz, S., Phillips, M. C., Segrest, J. P., and Anantharamaiah, G. M. (1995) *J. Biol. Chem.* **270**, 1602–1611
202. Casey, P. J. (1995) *Science* **268**, 221–225
203. McIlhinney, R. A. J. (1990) *Trends Biochem. Sci.* **15**, 387–391
204. Buss, J. E., Mumby, S. M., Casey, P. J., Gilman, A. G., and Sefton, B. M. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7493–7497
205. Johnson, D. R., Bhatnagar, R. S., Knoll, L. J., and Gordon, J. I. (1994) *Ann. Rev. Biochem.* **63**, 869–914
206. Peseckis, S. M., and Resh, M. D. (1994) *J. Biol. Chem.* **269**, 30888–30892
207. McLaughlin, S., and Aderem, A. (1995) *Trends Biochem. Sci.* **20**, 272–276
208. Milligan, G., Parenti, M., and Magee, A. I. (1995) *Trends Biochem. Sci.* **20**, 181–186
209. Wilson, P. T., and Bourne, H. R. (1995) *J. Biol. Chem.* **270**, 9667–9675
210. Duncan, J. A., and Gilman, A. G. (1996) *J. Biol. Chem.* **271**, 23594–23600
211. Marshall, C. J. (1993) *Science* **259**, 1865–1866
212. Glomset, J. A., Gelb, M. H., and Farnsworth, C. C. (1990) *Trends Biochem. Sci.* **15**, 139–142
213. Casey, P. J., and Seabra, M. C. (1996) *J. Biol. Chem.* **271**, 5289–5292
214. Shipton, C. A., Parmryd, I., Swiezewska, E., Andersson, B., and Dallner, G. (1995) *J. Biol. Chem.* **270**, 566–572
215. Wedegaertner, P. B., Wilson, P. T., and Bourne, H. R. (1995) *J. Biol. Chem.* **270**, 503–506
216. Kikuchi, A., Sagami, H., and Ogura, K. (1999) *J. Biol. Chem.* **274**, 18011–18016
217. Turner, A. J. (1994) in *Essays in Biochemistry*, Vol. 28 (Tipton, K. F., ed), pp. 113–128, Portland Press, London
218. Low, M. G. (1987) *Biochem. J.* **244**, 1–13
219. Lisanti, M. P., and Rodriguez-Boulan, E. (1990) *Trends Biochem. Sci.* **15**, 113–118
220. Thomas, J. R., Dwek, R. A., and Rademacher, T. W. (1990) *Biochemistry* **29**, 5413–5422
221. Brewis, I. A., Ferguson, M. A. J., Mehlert, A., Turner, A. J., and Hooper, N. M. (1995) *J. Biol. Chem.* **270**, 22946–22956
222. Previato, J. O., Jones, C., Xavier, M. T., Wait, R., Travassos, L. R., Parodi, A. J., and Mendonça-Previato, L. (1995) *J. Biol. Chem.* **270, No 13**, 7241–7250
223. Heinz, D. W., Ryan, M., Smith, M. P., Weaver, L. H., Keana, J. F. W., and Griffith, O. H. (1996) *Biochemistry* **35**, 9496–9504
224. Williams, A. F., and Gagnon, J. (1982) *Science* **216**, 696–703
225. Mehlert, A., Richardson, J. M., and Ferguson, M. A. J. (1998) *J. Mol. Biol.* **277**, 379–392
225a. Eisenhaber, B., Bork, P., Yuan, Y., Löffler, G., and Eisenhaber, F. (2000) *Trends Biochem. Sci.* **25**, 340–341
226. Reynolds, J. A., and Trayer, H. (1971) *J. Biol. Chem.* **246**, 7337–7342
227. Pryde, J. G. (1986) *Trends Biochem. Sci.* **11**, 160–163
228. Bennett, V. (1985) *Ann. Rev. Biochem.* **54**, 273–304
229. Justice, J. M., Murtagh, J. J., Jr., Moss, J., and Vaughan, M. (1995) *J. Biol. Chem.* **270**, 17970–17976
230. Neville, D. M., Jr., and Glossmann, H. (1971) *J. Biol. Chem.* **246**, 6335–6338
231. Inaba, M., and Maede, Y. (1988) *J. Biol. Chem.* **263**, 17763–17771
232. Steck, T. L. (1974) *J. Cell Biol.* **62**, 1–19
233. Luna, E. J., Kidd, G. H., and Branton, D. (1979) *J. Biol. Chem.* **254**, 2526–2532
234. Viitala, J., and Järnefelt, J. (1985) *Trends Biochem. Sci.* **10**, 392–395
235. Marchesi, V. T., Furthmayr, H., and Tomita, M. (1976) *Ann. Rev. Biochem.* **45**, 667–698
236. Jay, D., and Cantley, L. (1986) *Ann. Rev. Biochem.* **55**, 511–538
237. Murthy, S. N. P., Wilson, J., Zhang, Y., and Lorand, L. (1994) *J. Biol. Chem.* **269**, 22907–22911
238. Tanner, M. J. A. (1996) *Nature (London)* **382**, 209–210
239. Müller-Berger, S., Karbach, D., König, J., Lepke, S., Wood, P. G., Appelhans, H., and Passow, H. (1995) *Biochemistry* **34**, 9315–9324
240. Erickson, H. K. (1997) *Biochemistry* **36**, 9958–9967
241. Fukuda, M., Lauffenburger, M., Sasaki, H., Rogers, M. E., and Dell, A. (1987) *J. Biol. Chem.* **262**, 11952–11957
242. MacKenzie, K. R., Prestegard, J. H., and Engelman, D. M. (1997) *Science* **276**, 131–133
243. Challou, N., Goormaghtigh, E., Cabiaux, V., Conrath, K., and Ruysschaert, J.-M. (1994) *Biochemistry* **33**, 6902–6910
244. Huang, C.-H., Reid, M., Daniels, G., and Blumenfeld, O. O. (1993) *J. Biol. Chem.* **268**, 25902–25908
245. Marchesi, V. T. (1985) *Ann. Rev. Cell Biol.* **1**, 531–561
246. Fujita, T., Ralston, G. B., and Morris, M. B. (1998) *Biochemistry* **37**, 264–271
247. Yan, Y., Winograd, E., Viel, A., Cronin, T., Harrison, S. C., and Branton, D. (1993) *Science* **262**, 2027–2030
248. Cherry, L., Menhart, N., and Fung, L. W.-M. (1999) *J. Biol. Chem.* **274**, 2077–2084
249. Begg, G. E., Harper, S. L., Morris, M. B., and Speicher, D. W. (2000) *J. Biol. Chem.* **275**, 3279–3287
250. Fujita, T., Ralston, G. B., and Morris, M. B. (1998) *Biochemistry* **37**, 272–280
250a. Rief, M., Pascual, J., Saraste, M., and Gaub, H. E. (1999) *J. Mol. Biol.* **286**, 553–561
251. Fowler, V. M. (1986) *Nature (London)* **322**, 777–778
252. Carraway, K. L. (1975) *Biochim. Biophys. Acta.* **415**, 379–410
253. Byers, T. J., and Branton, D. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6153–6157
254. Elgsaeter, A., Stokke, B. T., Mikkelsen, A., and Branton, D. (1986) *Science* **234**, 1217–1223
255. Pantazatos, D. P., and MacDonald, R. I. (1997) *J. Biol. Chem.* **272**, 21052–21059
256. Luna, E. J., and Hitt, A. L. (1992) *Science* **258**, 955–964
257. Reichstein, E., and Blostein, R. (1975) *J. Biol. Chem.* **250**, 6256–6263
258. Bennett, V. (1992) *J. Biol. Chem.* **267**, 8703–8706
259. Michaely, P., and Bennett, V. (1993) *J. Biol. Chem.* **268**, 22703–22709
260. Kordeli, E., Lambert, S., and Bennett, V. (1995) *J. Biol. Chem.* **270**, 2352–2359
261. An, X.-L., Takakuwa, Y., Nunomura, W., Manno, S., and Mohandas, N. (1996) *J. Biol. Chem.* **271**, 33187–33191
262. Workman, R. F., and Low, P. S. (1998) *J. Biol. Chem.* **273**, 6171–6176
263. Malik, S., Sami, M., and Watts, A. (1993) *Biochemistry* **32**, 10078–10084
264. Azim, A. C., Knoll, J. H. M., Beggs, A. H., and Chishti, A. H. (1995) *J. Biol. Chem.* **270**, 17407–17413
265. Wolfe, L. C., John, K. M., Falcone, J. C., Byrne, A. M., and Lux, S. E. (1982) *N. Engl. J. Med.* **307**, 1367–1374
266. Diakowski, W., and Sikorski, A. F. (1995) *Biochemistry* **34**, 13252–13258
267. Marfatia, S. M., Lue, R. A., Branton, D., and Chishti, A. H. (1995) *J. Biol. Chem.* **270**, 715–719
268. Krauss, S. W., Chasis, J. A., Rogers, C., Mohandas, N., Krockmalnic, G., and Penman, S. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7297–7302
269. Khodadad, J. K., Waugh, R. E., Podolski, J. L., Josephs, R., and Steck, T. L. (1996) *Biophys. J.* **70**, 1036–1044
270. Hansen, J. C., Skalak, R., Chien, S., and Hoger, A. (1996) *Biophys. J.* **70**, 146–166
271. Rice-Evans, C. A., and Dunn, M. J. (1982) *Trends Biochem. Sci.* **7**, 282–286
272. Mohandas, N., and Evans, E. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 787–818
273. Agre, P., Casella, J. F., Zinkham, W. H., McMillan, C., and Bennett, V. (1985) *Nature (London)* **314**, 380–383
274. Baines, A. J., and Bennett, V. (1985) *Nature (London)* **315**, 410–413
275. Clark, E. A., and Brugge, J. S. (1995) *Science* **268**, 233–239
276. Dejana, E., Corada, M., and Lampugnani, M. G. (1995) *FASEB J.* **9**, 910–918
277. Malik, R. K., and Parsons, J. T. (1996) *J. Biol. Chem.* **271**, 29785–29791
278. Lo, S. H., An, Q., Bao, S., Wong, W.-K., Liu, Y., Janmey, P. A., Hartwig, J. H., and Chen, L. B. (1994) *J. Biol. Chem.* **269**, 22310–22319
279. Horwitz, A. F. (1997) *Sci. Am.* **276**(May), 68–75
280. de Pereda, J. M., Wiche, G., and Liddington, R. C. (1999) *EMBO J.* **18**, 4087–4095
281. Springer, T. A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 65–72
282. Bazzoni, G., and Hemler, M. E. (1998) *Trends Biochem. Sci.* **23**, 30–34
283. Emsley, J., King, S. L., Bergelson, J. M., and Liddington, R. C. (1997) *J. Biol. Chem.* **272**, 28512–28517
284. Luscinskas, F. W., and Lawler, J. (1994) *FASEB J.* **8**, 929–938
285. Simon, K. O., Otey, C. A., Pavelko, F. M., and Burridge, K. (1991) *Curr. Top. Membr. Transp.* **38**, 57–64
286. Edelman, G. M. (1983) *Science* **219**, 450–457
287. Becker, J. W., Erickson, H. P., Hoffman, S., Cunningham, B. A., and Edelman, G. M. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1088–1092
288. Cunningham, B. A., Hemperly, J. J., Murray, B. A., Prediger, E. A., Brackenbury, R., and Edelman, G. M. (1987) *Science* **236**, 799–806
289. Kudo, M., Kitajima, K., Inoue, S., Shiokawa, K., Morris, H. R., Dell, A., and Inoue, Y. (1996) *J. Biol. Chem.* **271**, 32667–32677
290. Newman, P. J., Berndt, M. C., Gorski, J., White, G. C., II, Lyman, S., Paddock, C., and Muller, W. A. (1990) *Science* **247**, 1219–1222
291. DeLisser, H. M., Yan, H. C., Newman, P. J., Muller, W. A., Buck, C. A., and Albelda, S. M. (1993) *J. Biol. Chem.* **268**, 16037–16046
291a. Chen, H. A., Pfuhl, M., McAlister, M. S. B., and Driscoll, P. C. (2000) *Biochemistry* **39**, 6814–6824
292. Casasnovas, J. M., Stehle, T., Liu, J.-h, Wang, J.-h, and Springer, T. A. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4134–4139
293. Wang, J.-H., Pepinsky, R. B., Stehle, T., Liu, J.-H., Karpusas, M., Browning, B., and Osborn, L. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5714–5718
294. Springer, T. A. (1990) *Nature (London)* **346**, 425–434

## References


295. Wagner, G. (1995) *Science* **267**, 342
296. Nagar, B., Overduin, M., Ikura, M., and Rini, J. M. (1996) *Nature (London)* **380**, 360–364
297. Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehmann, M. S., Grübel, G., Legrand, J.-F., Als-Nielsen, J., Colman, D. R., and Hendrickson, W. A. (1995) *Nature (London)* **374**, 327–336
298. Overduin, M., Harvey, T. S., Bagby, S., Tong, K. I., Yau, P., Takeichi, M., and Ikura, M. (1995) *Science* **267**, 386–389

298a. Nollet, F., Kools, P., and van Roy, F. (2000) *J. Mol. Biol.* **299**, 551–572

298b. Pertz, O., Bozic, D., Koch, A. W., Fauser, C., Brancaccio, A., and Engel, J. (1999) *EMBO J.* **18**, 1738–1747

299. Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M., and Rossmann, M. G. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4140–4145
300. Patel, D. J., and Gumbiner, B. M. (1995) *Nature (London)* **374**, 306–307
301. Berx, G., Cleton-Jansen, A.-M., Nollet, F., de Leeuw, W. J. F., van de Vijver, M. J., Cornelisse, C., and van Roy, F. (1995) *EMBO J.* **14**, 6107–6115
302. Vartio, T., and Vaheri, A. (1983) *Trends Biochem. Sci.* **8**, 442–444
303. Hynes, R. O. (1986) *Sci. Am.* **254**(Jun), 42–51
304. Ruoslahti, E. (1988) *Ann. Rev. Biochem.* **57**, 375–413
305. Engel, J., Odermatt, E., Engel, A., Madri, J. A., Furthmayr, H., Rohde, H., and Timpl, R. (1981) *J. Mol. Biol.* **150**, 97–120
306. Ingham, K. C., Brew, S. A., Huff, S., and Litvinovich, S. V. (1997) *J. Biol. Chem.* **272**, 1718–1724
307. Johnson, K. J., Sage, H., Briscoe, G., and Erickson, H. P. (1999) *J. Biol. Chem.* **274**, 15473–15479
308. Grant, R. P., Spitzfaden, C., Altroff, H., Campbell, I. D., and Mardon, H. J. (1997) *J. Biol. Chem.* **272**, 6159–6166
309. Hynes, R. O. (1990) *Fibronectins*, Springer-Verlag, New York
310. Moyano, J. V., Carnemolla, B., Domínguez-Jiménez, C., García-Gila, M., Albar, J. P., Sánchez-Aparicio, P., Leprini, A., Querzé, G., Zardi, L., and Garcia-Pardo, A. (1997) *J. Biol. Chem.* **272**, 24832–24836
311. Greenberg, C. S., Birckbichler, P. J., and Rice, R. H. (1991) *FASEB J.* **5**, 3071–3077
312. Jeong, J.-M., Murthy, S. N. P., Radek, J. T., and Lorand, L. (1995) *J. Biol. Chem.* **270**, 5654–5658
313. An, S. S. A., Jiménez-Barbero, J., Petersen, T. E., and Llinás, M. (1992) *Biochemistry* **31**, 9927–9933
314. Spitzfaden, C., Grant, R. P., Mardon, H. J., and Campbell, L. D. (1997) *J. Mol. Biol.* **265**, 565–579
315. Kornblihtt, A. R., Pesce, C. G., Alonso, C. R., Cramer, P., Srebrow, A., Werbajh, S., and Muro, A. F. (1996) *FASEB J.* **10**, 248–257
316. Yoneda, A., Ogawa, H., Kojima, K., and Matsumoto, I. (1998) *Biochemistry* **37**, 6351–6360
317. Fischer, D., Chiquet-Ehrismann, R., Bernasconi, C., and Chiquet, M. (1995) *J. Biol. Chem.* **270**, 3378–3384
318. Yokosaki, Y., Matsuura, N., Higashiyama, S., Murakami, I., Obara, M., Yamakido, M., Shigeto, N., Chen, J., and Sheppard, D. (1998) *J. Biol. Chem.* **273**, 11423–11428
319. Denda, S., Müller, U., Crossin, K. L., Erickson, H. P., and Reichardt, L. F. (1998) *Biochemistry* **37**, 5464–5474
320. Oberhauser, A. F., Marszalek, P. E., Erickson, H. P., and Fernandez, J. M. (1998) *Nature (London)* **393**, 181–185
321. Stetefeld, J., Mayer, U., Timpl, R., and Huber, R. (1996) *J. Mol. Biol.* **257**, 644–657
322. Colognato, H., MacCarrick, M., O'Rear, J. J., and Yurchenco, P. D. (1997) *J. Biol. Chem.* **272**, 29330–29336
323. Lane, T. F., and Sage, E. H. (1994) *FASEB J.* **8**, 163–173
324. Venstrom, K. A., and Reichardt, L. F. (1993) *FASEB J.* **7**, 996–1003
325. Gherzi, R., Carnemolla, B., Siri, A., Ponassi, M., Balza, E., and Zardi, L. (1995) *J. Biol. Chem.* **270**, 3429–3434
326. Ruoslahti, E., and Pierschbacher, M. D. (1987) *Science* **238**, 491–497
327. D'Souza, S. E., Ginsberg, M. H., and Plow, E. F. (1991) *Trends Biochem. Sci.* **16**, 246–250
328. Leahy, D. J., Hendrickson, W. A., Aukhil, I., and Erickson, H. P. (1992) *Science* **258**, 987–991
329. Stanley, K. K., Newby, A. C., and Luzio, J. P. (1982) *Trends Biochem. Sci.* **7**, 145–147
330. Roberts, M. F. (1996) *FASEB J.* **10**, 1159–1172
331. Han, S. K., Yoon, E. T., Scott, D. L., Sigler, P. B., and Cho, W. (1997) *J. Biol. Chem.* **272**, 3573–3582
332. Caruthers, A., and Melchior, D. L. (1986) *Trends Biochem. Sci.* **11**, 331–335
333. Gutknecht, J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6443–6446
334. Paula, S., Volkov, A. G., and Deamer, D. W. (1998) *Biophys. J.* **74**, 319–327
335. Stein, W. D. (1990) *Channels, Carriers and Pumps: an Introduction to Membrane Transport*, Academic Press, San Diego, California

335a. Paulsen, I. T., Nguyen, L., Sliwinski, M. K., Rabus, R., and Saier, M. H., Jr. (2000) *J. Mol. Biol.* **301**, 75–100

336. Fox, R. O., Jr., and Richards, F. M. (1982) *Nature (London)* **300**, 325–330
337. Kempf, C., Klausner, R. D., Weinstein, J. N., Van Renswoude, J., Pincus, M., and Blumenthal, R. (1982) *J. Biol. Chem.* **257**, 2469–2476
338. Klebba, P. E., Hofnung, M., and Charbit, A. (1994) *EMBO J.* **13**, 4670–4675
339. Schirmer, T., Keller, T. A., Wang, Y.-F., and Rosenbusch, J. P. (1995) *Science* **267**, 512–514
340. Meyer, J. E. W., Hofnung, M., and Schulz, G. E. (1997) *J. Mol. Biol.* **266**, 761–775
341. Jordy, M., Andersen, C., Schülein, K., Ferenci, T., and Benz, R. (1996) *J. Mol. Biol.* **259**, 666–678

341a. Dumas, F., Koebnik, R., Winterhalter, M., and Van Gelder, P. (2000) *J. Biol. Chem.* **275**, 19747–19751

342. Samartzidou, H., and Delcour, A. H. (1998) *EMBO J.* **17**, 93–100
343. Pauptit, R. A., Schirmer, T., Jansonius, J. N., Rosenbusch, J. P., Parker, M. W., Tucker, A. D., Tsernoglou, D., Weiss, M. S., and Schulz, G. E. (1991) *J. Structural Biol.* **107**, 136–145
344. Kreusch, A., Neubüser, A., Schiltz, E., Weckesser, J., and Schulz, G. E. (1994) *Protein Sci.* **3**, 58–63

344a. Müller, D. J., and Engel, A. (1999) *J. Mol. Biol.* **285**, 1347–1351

345. Bainbridge, G., Mobasheri, H., Armstrong, G. A., Lea, E. J. A., and Lakey, J. H. (1998) *J. Mol. Biol.* **275**, 171–176
346. Van Gelder, P., Saint, N., Phale, P., Eppens, E. F., Prilipov, A., van Boxtel, R., Rosenbusch, J. P., and Tommassen, J. (1997) *J. Mol. Biol.* **269**, 468–472
347. Lyer, R., Wu, Z., Woster, P. M., and Delcour, A. H. (2000) *J. Mol. Biol.* **297**, 933–945
348. Watanabe, M., Rosenbusch, J., Schirmer, T., and Karplus, M. (1997) *Biophys. J.* **72**, 2094–2102
349. Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Jansonius, J. N., and Rosenbusch, J. P. (1992) *Nature (London)* **358**, 727–733
350. Stathopoulos, C. (1996) *Protein Sci.* **5**, 170–173
351. Mannella, C. A. (1992) *Trends Biochem. Sci.* **17**, 315–320
352. Li, A. Z., Huang, H., Re, X., Qi, L. J., and Marx, K. A. (1998) *Biophys. J.* **74**, 964–973
353. Reumann, S., Maier, E., Benz, R., and Heldt, H. W. (1995) *J. Biol. Chem.* **270**, 17559–17565
354. Chrispeels, M. J., and Agre, P. (1994) *Trends Biochem. Sci.* **19**, 421–425
355. Knepper, M. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6255–6258

355a. Borgnia, M., Nielsen, S., Engel, A., and Agre, P. (1999) *Ann. Rev. Biochem.* **68**, 425–458

356. Ringler, P., Borgnia, M. J., Stahlberg, H., Maloney, P. C., Agre, P., and Engel, A. (1999) *J. Mol. Biol.* **291**, 1181–1190
357. de Groot, B. L., Heymann, J. B., Engel, A., Mitsuoka, K., Fujiyoshi, Y., and Grubmüller, H. (2000) *J. Mol. Biol.* **300**, 987–994
358. Cheng, A., van Hoek, A. N., Yeager, M., Verkman, A. S., and Mitra, A. K. (1997) *Nature (London)* **387**, 627–630
359. Ma, T., Song, Y., Gillespie, A., Carlson, E. J., Epstein, C. J., and Verkman, A. S. (1999) *J. Biol. Chem.* **274**, 20071–20074

359a. Zeuthen, T., and Klaerke, D. A. (1999) *J. Biol. Chem.* **274**, 21631–21636

360. Barone, L. M., Shih, C., and Wasserman, B. P. (1997) *J. Biol. Chem.* **272**, 30672–30677
361. Calamita, G., Bishai, W. R., Preston, G. M., Guggino, W. B., and Agre, P. (1995) *J. Biol. Chem.* **270**, 29063–29066
362. Deen, P. M. T., Verdijk, M. A. J., Knoers, N. V. A. M., Wieringa, B., Monnens, L. A. H., van OS, C. H., and van Oost, B. A. (1994) *Science* **264**, 92–95
363. Bai, L., Fushimi, K., Sasaki, S., and Marumo, F. (1996) *J. Biol. Chem.* **271**, 5171–5176
364. Jan, L. Y., and Jan, Y. N. (1994) *Nature (London)* **371**, 119–122
365. Armstrong, C. (1998) *Science* **280**, 56–57
366. Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., and MacKinnon, R. (1998) *Science* **280**, 69–77
367. Heginbothan, L., Odessey, E., and Miller, C. (1997) *Biochemistry* **36**, 10335–10342
368. MacKinnon, R., Cohen, S. L., Kuo, A., Lee, A., and Chait, B. T. (1998) *Science* **280**, 106–109
369. Aidley, D. J., and Stanfield, P. R. (1996) *Ion Channels, Molecules in Action*, Cambridge Univ. Press, New York
370. Kreusch, A., Pfaffinger, P. J., Stevens, C. F., and Choe, S. (1998) *Nature (London)* **392**, 945–948

370a. Gulbis, J. M., Zhou, M., Mann, S., and MacKinnon, R. (2000) *Science* **289**, 123–127

370b. Kobertz, W. R., Williams, C., and Miller, C (2000) *Biochemistry* **39**, 10347–10352

370c. Roux, B., Bernèche, S., and Im, W. (2000) *Biochemistry* **39**,

371. Klingenberg, M. (1981) *Nature (London)* **290**, 449–454
372. Prince, R. C., Gunson, D. E., and Scarpa, A. (1985) *Trends Biochem. Sci.* **10**, 99
373. Smith, R., Separovic, F., Milne, T. J., Whittaker, A., Bennett, F. M., Cornell, B. A., and Makriyannis, A. (1994) *J. Mol. Biol.* **241**, 456–466
374. Dathe, M., Schümann, M., Wieprecht, T., Winkler, A., Beyermann, M., Krause, E., Matsuzaki, K., Murase, O., and Bienert, M. (1996) *Biochemistry* **35**, 12612–12622
375. Kiyota, T., Lee, S., and Sugihara, G. (1996) *Biochemistry* **35**, 13196–13204
376. Monette, M., and Lafleur, M. (1996) *Biophys. J.* **70**, 2195–2202
377. Ojcius, D. M., and Young, J. D.-E. (1991) *Trends Biochem. Sci.* **16**, 225–229
378. Langs, D. A. (1988) *Science* **241**, 188–191

## References


379. Doyle, D. A., and Wallace, B. A. (1997) *J. Mol. Biol.* **266**, 963–977
380. Iqbal, M., and Balaram, P. (1981) *Biochemistry* **20**, 7278–7284
381. Malashkevich, V. N., Kammerer, R. A., Efimov, V. P., Schulthess, T., and Engel, J. (1996) *Science* **274**, 761–765
382. Unwin, N. (1995) *Nature (London)* **373**, 37–43
383. Harrold, F. M., and Baarda, J. R. (1967) *J. Bacteriol.* **94**, 53–60
384. Harold, F. M., and Van Brunt, J. (1977) *Science* **197**, 372–373
385. Pressman, B. C. (1976) *Ann. Rev. Biochem.* **45**, 501–530
386. Hamilton, J. A., Sabesan, M. N., and Steinrauf, L. K. (1981) *J. Am. Chem. Soc.* **103**, 5880–5885
387. Duax, W. L., Hauptman, M., Weeks, C. M., and Norton, D. A. (1972) *Science* **176**, 911–914
388. Karle, I. L., Karle, J., Wieland, T., Burgemeister, W., Faulstich, H., and Witkop, B. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 1836–1840
389. Pangborn, W., Duax, W., and Langs, D. (1987) *J. Am. Chem. Soc.* **109**, 2163–2165
390. Karle, I. L. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7155–7159
391. Inabayashi, M., Miyauchi, S., Kamo, N., and Jin, T. (1995) *Biochemistry* **34**, 3455–3460
392. Ledger, P. W., and Tanzer, M. L. (1984) *Trends Biochem. Sci.* **9**, 313–314
393. Marger, M. D., and Saier, M. H., Jr. (1993) *Trends Biochem. Sci.* **18**, 13–20
394. Goswitz, V. C., and Brooker, R. J. (1995) *Protein Sci.* **4**, 534–537
395. Seidel, H. M., and Knowles, J. R. (1994) *Biochemistry* **33**, 5641–5646
396. Cloherty, E. K., Sulzman, L. A., Zottola, R. J., and Carruthers, A. (1995) *Biochemistry* **34**, 15395–15406
397. Shanahan, M. F., and D'Artel-Ellis, J. (1984) *J. Biol. Chem.* **259**, 13878–13884
398. Bell, G. I., Burant, C. F., Takeda, J., and Gould, G. W. (1993) *J. Biol. Chem.* **268**, 19161–19164
399. Silverman, M. (1991) *Ann. Rev. Biochem.* **60**, 757–794
400. Mueckler, M. (1994) *Eur. J. Biochem.* **219**, 713–725
401. Maher, F., Vannucci, S. J., and Simpson, I. A. (1994) *FASEB J.* **8**, 1003–1011
402. Burant, C. F., Takeda, J., Brot-Laroche, E., Bell, G. I., and Davidson, N. O. (1992) *J. Biol. Chem.* **267**, 14523–14526
403. Zamora-León, S. P., Golde, D. W., Concha, I. I., Rivas, C. I., Delgado-López, F., Baselga, J., Nualart, F., and Vera, J. C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1847–1852
404. Coderre, L., Kandror, K. V., Vallega, G., and Pilch, P. F. (1995) *J. Biol. Chem.* **270**, 27584–27588
405. Fischer, Y., Thomas, J., Sevilla, L., Munoz, P., Becker, C., Holman, G., Kozka, I. J., Palacín, M., Testar, X., Kammermeier, H., and Zorzano, A. (1997) *J. Biol. Chem.* **272**, 7085–7092

405a. Shepherd, P. R., and Kahn, B. B. (1999) *N. Engl. J. Med.* **341**, 248–257

406. Mueckler, M., and Makepeace, C. (1997) *J. Biol. Chem.* **272**, 30141–30146
407. Cloherty, E. K., Diamond, D. L., Heard, K. S., and Carruthers, A. (1996) *Biochemistry* **35**, 13231–13239
408. Seatter, M. J., De La Rue, S. A., Porter, L. M., and Gould, G. W. (1998) *Biochemistry* **37**, 1322–1326
409. Rumsey, S. C., Kwon, O., Xu, G. W., Burant, C. F., Simpson, I., and Levine, M. (1997) *J. Biol. Chem.* **272**, 18982–18989
410. Wiese, T. J., Dunlap, J. A., and Yorek, M. A. (1994) *J. Biol. Chem.* **269**, 22705–22711
411. Kasahara, M., Shimoda, E., and Maeda, M. (1997) *J. Biol. Chem.* **272**, 16721–16724
412. Turk, E., Kerner, C. J., Lostao, M. P., and Wright, E. M. (1996) *J. Biol. Chem.* **271**, 1925–1934
413. Chen, X.-Z., Coady, M. J., Jackson, F., Berteloot, A., and Lapointe, J.-Y. (1995) *Biophys. J.* **69**, 2405–2414
414. Panayotova-Heiermann, M., Eskandari, S., Turk, E., Zampighi, G. A., and Wright, E. M. (1997) *J. Biol. Chem.* **272**, 20324–20327
415. Prasad, P. D., Wang, H., Kekuda, R., Fujita, T., Fei, Y.-J., Devoe, L. D., Leibach, F. H., and Ganapathy, V. (1998) *J. Biol. Chem.* **273**, 7501–7506
416. Malandro, M. S., and Kilberg, M. S. (1996) *Ann. Rev. Biochem.* **65**, 305–336
417. Utsunomiya-Tate, N., Endou, H., and Kanai, Y. (1996) *J. Biol. Chem.* **271**, 14883–14890
418. Kong, C.-T., Yet, S.-F., and Lever, J. E. (1993) *J. Biol. Chem.* **268**, 1509–1512
419. Su, T.-Z., Wang, M., Syu, L.-J., Saltiel, A. R., and Oxender, D. L. (1998) *J. Biol. Chem.* **273**, 3173–3179
420. Yao, D., Mackenzie, B., Ming, H., Varoqui, H., Zhu, H., Hediger, M. A., and Erickson, J. D. (2000) *J. Biol. Chem.* **275**, 22790–22797

420a. Reimer, R. J., Chaudhry, F. A., Gray, A. T., and Edwards, R. H. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7715–7720

421. Mosckovitz, R., Udenfriend, S., Felix, A., Heimer, E., and Tate, S. S. (1994) *FASEB J.* **8**, 1069–1074
422. Chillarón, J., Estévez, R., Mora, C., Wagner, C. A., Suessbrich, H., Lang, F., Gelpí, J. L., Testar, X., Busch, A. E., Zorzano, A., and Palacín, M. (1996) *J. Biol. Chem.* **271**, 17761–17770
423. Trotti, D., Rossi, D., Gjesdal, O., Levy, L. M., Racagni, G., Danbolt, N. C., and Volterra, A. (1996) *J. Biol. Chem.* **271**, 5976–5979
424. Peghini, P., Janzen, J., and Stoffel, W. (1997) *EMBO J.* **16**, 3822–3832
425. Lebrun, B., Sakaitani, M., Shimamoto, K., Yasuda-Kamatani, Y., and Nakajima, T. (1997) *J. Biol. Chem.* **272**, 20336–20339
426. Ponce, J., Biton, B., Benavides, J., Avenet, P., and Aragón, C. (2000) *J. Biol. Chem.* **275**, 13856–13862
427. Pantanowitz, S., Bendahan, A., and Kanner, B. I. (1993) *J. Biol. Chem.* **268**, 3222–3225
428. Bennett, E. R., and Kanner, B. I. (1997) *J. Biol. Chem.* **272**, 1203–1210
429. Velaz-Faircloth, M., Guadano-Ferraz, A., Henzi, V. A., and Fremeau, R. T., Jr. (1995) *J. Biol. Chem.* **270**, 15755–15761
430. Chen, J.-G., Liu-Chen, S., and Rudnick, G. (1998) *J. Biol. Chem.* **273**, 12675–12681
431. Nikaido, H., and Saier, M. H., Jr. (1992) *Science* **258**, 936–942
432. Will, A., Grassl, R., Erdmenger, J., Caspari, T., and Tanner, W. (1998) *J. Biol. Chem.* **273**, 11456–11462
433. Kaback, H. R., Bibi, E., and Roepe, P. D. (1990) *Trends Biochem. Sci.* **15**, 309–314
434. Frillingos, S., Wu, J., Venkatesan, P., and Kaback, H. R. (1997) *Biochemistry* **36**, 6408–6414
435. Green, A. L., Anderson, E. J., and Brooker, R. J. (2000) *J. Biol. Chem.* **275**, 23240–23246
436. Zhao, M., Zen, K.-C., Hubbell, W. L., and Kaback, H. R. (1999) *Biochemistry* **38**, 7407–7412
437. Zani, M. L., Pourcher, T., and Leblanc, G. (1994) *J. Biol. Chem.* **269**, 24883–24889
438. Pourcher, T., Bibi, E., Kaback, H. R., and Leblanc, G. (1996) *Biochemistry* **35**, 4161–4168
439. McDonald, T. P., Walmsley, A. R., and Henderson, P. J. F. (1997) *J. Biol. Chem.* **272**, 15189–15199
440. Fei, Y.-J., Liu, W., Prasad, P. D., Kekuda, R., Oblak, T. G., Ganapathy, V., and Leibach, F. H. (1997) *Biochemistry* **36**, 452–460
441. Brandsch, M., Thunecke, F., Küllertz, G., Schutkowski, M., Fischer, G., and Neubert, K. (1998) *J. Biol. Chem.* **273**, 3861–3864
442. Boorer, K. J., Frommer, W. B., Bush, D. R., Kreman, M., Loo, D. D. F., and Wright, E. M. (1996) *J. Biol. Chem.* **271**, 2213–2220
443. Boorer, K. J., and Fischer, W.-N. (1997) *J. Biol. Chem.* **272**, 13040–13046
444. Someya, Y., Niwa, A., Sawai, T., and Yamaguchi, A. (1995) *Biochemistry* **34**, 7–12
445. Nikaido, K., Liu, P.-Q., and Ames, G. F.-L. (1997) *J. Biol. Chem.* **272**, 27745–27752
446. Létoffé, S., Delepelaire, P., and Wandersman, C. (1996) *EMBO J.* **15**, 5804–5811
447. Blattner, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science* **277**, 1453–1462
448. Decottignies, A., Lambert, L., Catty, P., Degand, H., Epping, E. A., Moye-Rowley, W. S., Balzi, E., and Goffeau, A. (1995) *J. Biol. Chem.* **270**, 18150–18157

448a. Driessen, A. J. M., Rosen, B. P., and Konings, W. N. (2000) *Trends Biochem. Sci.* **25**, 397–401

449. Sharom, F. J., DiDiodato, G., Yu, X., and Ashbourne, K. J. D. (1995) *J. Biol. Chem.* **270**, 10334–10341
450. Ko, Y. H., and Pedersen, P. L. (1995) *J. Biol. Chem.* **270**, 22093–22096
451. Nossal, N. G., and Heppel, L. A. (1966) *J. Biol. Chem.* **241**, 3055–3062
452. Ames, G. F.-L. (1986) *Ann. Rev. Biochem.* **55**, 397–425
453. Ames, G. F.-L., and Higgins, C. F. (1983) *Trends Biochem. Sci.* **8**, 97–100
454. Yao, N., Ledvina, P. S., Choudhary, A., and Quiocho, F. A. (1996) *Biochemistry* **35**, 2079–2085
455. Quiocho, F. A., and Vyas, N. K. (1984) *Nature (London)* **310**, 381–386
456. Gilliland, G. L., and Quiocho, F. A. (1981) *J. Mol. Biol.* **146**, 341–362
457. Vyas, N. K., Vyas, M. N., and Quiocho, F. A. (1988) *Science* **242**, 1290–1295
458. Vyas, N. K., Vyas, M. N., and Quiocho, F. A. (1991) *J. Biol. Chem.* **266**, 5226–5237
459. Shilton, B. H., Shuman, H. A., and Mowbray, S. L. (1996) *J. Mol. Biol.* **264**, 364–376
460. Oh, B.-H., Kang, C.-H., De Bondt, H., Kim, S.-H., Nikaido, K., Joshi, A. K., and Ames, G. F.-L. (1994) *J. Biol. Chem.* **269**, 4135–4143
461. Oh, B.-H., Ames, G. F.-L., and Kim, S.-H. (1994) *J. Biol. Chem.* **269**, 26323–26330
462. Sugiyama, S., Vassylyev, D. G., Matsushima, M., Kashiwagi, K., Igarashi, K., and Morikawa, K. (1996) *J. Biol. Chem.* **271**, 9519–9525
463. Pflugrath, J. W., and Quiocho, F. A. (1988) *J. Mol. Biol.* **200**, 163–180
464. Manson, M. D., Boos, W., Bassford, P. J., Jr., and Rasmussen, B. A. (1985) *J. Biol. Chem.* **260**, 9727–9733
465. Reizer, J., Michotey, V., Reizer, A., and Saier, M. H., Jr. (1994) *Protein Sci.* **3**, 440–450
466. Garrett, D. S., Seok, Y.-J., Liao, D.-I., Peterkofsky, A., Gronenborn, A. M., and Clore, G. M. (1997) *Biochemistry* **36**, 2517–2530
467. Chauvin, F., Brand, L., and Roseman, S. (1994) *J. Biol. Chem.* **269**, 20263–20269

467a. Zhu, P.-P., Szczepanowski, R. H., Nosworthy, N. J., Ginsburg, A., and Peterkofsky, A. (1999) *Biochemistry* **38**, 15470–15479

468. Jones, B. E., Rajagopal, P., and Klevit, R. E. (1997) *Protein Sci.* **6**, 2107–2119
469. El-Kabbani, O. A. L., Waygood, E. B., and Delbaera, L. T. J. (1987) *J. Biol. Chem.* **262**, 12926–12929

## References


470. Reizer, J., Hoischen, C., Reizer, A., Pham, T. N., and Saier, M. H., Jr. (1993) *Protein Sci.* **2**, 506–521
471. Eberstadt, M., Grdadolnik, S. G., Gemmecker, G., Kessler, H., Buhr, A., and Erni, B. (1996) *Biochemistry* **35**, 11286–11292
472. Van Dijk, A. A., Scheek, R. M., Dijkstra, K., Wolters, G. K., and Robillard, G. T. (1992) *Biochemistry* **31**, 9063–9072
473. Gemmecker, G., Eberstadt, M., Buhr, A., Lanz, R., Golic Grdadolnik, S., Kessler, H., and Erni, B. (1997) *Biochemistry* **36**, 7408–7417
474. Pelton, J. G., Torchia, D. A., Remington, S. J., Murphy, K. P., Meadow, N. D., and Roseman, S. (1996) *J. Biol. Chem.* **271**, 33446–33456
475. Lanz, R., and Erni, B. (1998) *J. Biol. Chem.* **273**, 12239–12243
476. Nunn, R. S., Markovic-Housley, Z., Génovésio-Taverne, J.-C., Flükiger, K., Rizkallah, P. J., Jansonius, J. N., Schirmer, T., and Erni, B. (1996) *J. Mol. Biol.* **259**, 502–511
477. Roseman, S., and Meadow, N. D. (1990) *J. Biol. Chem.* **265**, 2993–2996
478. Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W., and Remington, S. J. (1993) *Science* **259**, 673–677
479. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry*, Freeman, San Francisco, California
480. Masuda, T., Dobson, G. P., and Veech, R. L. (1990) *J. Biol. Chem.* **265**, 20321–20334
481. Catterall, W. A. (1995) *Ann. Rev. Biochem.* **64**, 493–531
482. Valverde, M. A., Hardy, S. P., and Sepúlveda, F. V. (1995) *FASEB J.* **9**, 509–515
483. Schmidt-Rose, T., and Jentsch, T. J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7633–7638
484. Jentsch, T. J., Steinmeyer, K., and Schwarz, G. (1990) *Nature (London)* **348**, 510–514
485. Middleton, R. E., Pheasant, D. J., and Miller, C. (1994) *Biochemistry* **33**, 13189–13198
486. Schwappach, B., Stobrawa, S., Hechenberger, M., Steinmeyer, K., and Jentsch, T. J. (1998) *J. Biol. Chem.* **273**, 15110–15118
487. Gronemeier, M., Condie, A., Prosser, J., Steinmeyer, K., Jentsch, T. J., and Jockusch, H. (1994) *J. Biol. Chem.* **269**, 5963–5967
488. Pucéat, M., Korichneva, I., Cassoly, R., and Vassort, G. (1995) *J. Biol. Chem.* **270**, 1315–1322
489. Jarolim, P., Shayakul, C., Prabakaran, D., Jiang, L., Stuart-Tilley, A., Rubin, H. L., Simova, S., Zavadil, J., Herrin, J. T., Brouillette, J., Somers, M. J. G., Seemanova, E., Brugnara, C., Guay-Woodford, L. M., and Alper, S. L. (1998) *J. Biol. Chem.* **273**, 6380–6388

489a. Fujinaga, J., Tang, X.-B., and Casey, J. R. (1999) *J. Biol. Chem.* **274**, 6626–6633

490. Welsh, M. J., and Smith, A. E. (1995) *Sci. Am.* **273**(Dec), 52–59
491. Seibert, F. S., Linsdell, P., Loo, T. W., Hanrahan, J. W., Riordan, J. R., and Clarke, D. M. (1996) *J. Biol. Chem.* **271**, 27493–27499
492. Romero, M. F., Hediger, M. A., Boulpaep, E. L., and Boron, W. F. (1997) *Nature (London)* **387**, 409–413
493. Burnham, C. E., Amlal, H., Wang, Z., Shull, G. E., and Soleimani, M. (1997) *J. Biol. Chem.* **272**, 19111–19114
494. Muchhal, U. S., Pardo, J. M., and Raghothama, K. G. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10519–10523
495. Smith, F. W., Ealing, P. M., Hawkesford, M. J., and Clarkson, D. T. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9373–9377
496. Zhou, J.-J., Theodoulou, F. L., Muldin, I., Ingemarsson, B., and Miller, A. J. (1998) *J. Biol. Chem.* **273**, 12017–12023
497. Schroers, A., Burkovski, A., Wohlrab, H., and Krämer, R. (1998) *J. Biol. Chem.* **273**, 14269–14276
498. Phelps, A., Briggs, C., Mincone, L., and Wohlrab, H. (1996) *Biochemistry* **35**, 10757–10762
499. Collins, J. F., and Ghishan, F. K. (1996) *FASEB J.* **10**, 751–759
500. Garcia, C. K., Brown, M. S., Pathak, R. K., and Goldstein, J. L. (1995) *J. Biol. Chem.* **270**, 1843–1849
501. Kuo, S.-M., and Aronson, P. S. (1996) *J. Biol. Chem.* **271**, 15491–15497
502. Yellen, G. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 227–246
503. Miller, C. (1991) *Science* **252**, 1092–1096
504. Philipson, L. H., and Steiner, D. F. (1995) *Science* **268**, 372–373
505. Inagaki, N., Gonoi, T., Clement, J. P., IV, Namba, N., Inazawa, J., Gonzalez, G., Aguilar-Bryan, L., Seino, S., and Bryan, J. (1995) *Science* **270**, 1166–1170
506. Bränström, R., Efendic, S., Berggren, P.-O., and Larsson, O. (1998) *J. Biol. Chem.* **273**, 14113–14118
507. Ruknudin, A., Schulze, D. H., Sullivan, S. K., Lederer, W. J., and Welling, P. A. (1998) *J. Biol. Chem.* **273**, 14165–14171
508. Day, M. L., Johnson, M. H., and Cook, D. I. (1998) *EMBO J.* **17**, 1952–1960
509. McCormick, K. A., Isom, L. L., Ragsdale, D., Smith, D., Scheuer, T., and Catterall, W. A. (1998) *J. Biol. Chem.* **273**, 3954–3962
510. Favre, I., Moczydlowski, E., and Schild, L. (1996) *Biophys. J.* **71**, 3110–3125

510a. Lipkind, G. M., and Fozzard, H. A. (2000) *Biochemistry* **39**, 8161–8170

511. Pérez-García, M. T., Chiamvimonvat, N., Marban, E., and Tomaselli, G. F. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 300–304
512. Kosari, F., Sheng, S., Li, J., Mak, D.-O. D., Foskett, J. K., and Kleyman, T. R. (1998) *J. Biol. Chem.* **273**, 13469–13474
513. Garty, H. (1994) *FASEB J.* **8**, 522–528
514. Miller, R. J. (1992) *J. Biol. Chem.* **267**, 1403–1406
515. Mitterdorfer, J., Sinnegger, M. J., Grabner, M., Striessnig, J., and Glossmann, H. (1995) *Biochemistry* **34**, 9350–9355
516. McPherson, P. S., and Campbell, K. P. (1993) *J. Biol. Chem.* **268**, 13765–13768
517. Jeyakumar, L. H., Copello, J. A., O'Malley, A. M., Wu, G.-M., Grassucci, R., Wagenknecht, T., and Fleischer, S. (1998) *J. Biol. Chem.* **273**, 16011–16020
518. Kraus, R. L., Sinnegger, M. J., Glossmann, H., Hering, S., and Striessnig, J. (1998) *J. Biol. Chem.* **273**, 5586–5590
519. Kim, S., Lakhani, V., Costa, D. J., Sharara, A. I., Fitz, J. G., Huang, L.-W., Peters, K. G., and Kindman, L. A. (1995) *J. Biol. Chem.* **270**, 5266–5269
520. Mezna, M., and Michelangeli, F. (1996) *J. Biol. Chem.* **271**, 31818–31823
521. Nicoll, D. A., Hryshko, L. V., Matsuoka, S., Frank, J. S., and Philipson, K. D. (1996) *J. Biol. Chem.* **271**, 13385–13391
522. Crumpton, M. J., and Dedman, J. R. (1990) *Nature (London)* **345**, 212
523. Swairjo, M. A., and Seaton, B. A. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 193–213
524. Benz, J., Bergner, A., Hofmann, A., Demange, P., Göttig, P., Liemann, S., Huber, R., and Voges, D. (1996) *J. Mol. Biol.* **260**, 638–643
525. Burger, A., Berendes, R., Liemann, S., Benz, J., Hofmann, A., Göttig, P., Huber, R., Gerke, V., Thiel, C., Römisch, J., and Weber, K. (1996) *J. Mol. Biol.* **257**, 839–847
526. Demange, P., Voges, D., Benz, J., Liemann, S., Göttig, P., Berendes, R., Burger, A., and Huber, R. (1994) *Trends Biochem. Sci.* **19**, 272–276
527. Campos, B., Mo, Y. D., Mealy, T. R., Li, C. W., Swairjo, M. A., Balch, C., Head, J. F., Retzinger, G., Dedman, J. R., and Seaton, B. A. (1998) *Biochemistry* **37**, 8004–8010
528. Ling, G. N. (1984) *In Search of the Physical Basis of Life*, Plenum, New York
529. Baker, P. F., and Willis, J. S. (1970) *Nature (London)* **226**, 521–523
530. Kyte, J. (1972) *J. Biol. Chem.* **247**, 7642–7649
531. Lingrel, J. B., and Kuntzweiler, T. (1994) *J. Biol. Chem.* **269**, 19659–19662
532. Lutsenko, S., Daoud, S., and Kaplan, J. H. (1997) *J. Biol. Chem.* **272**, 5249–5255
533. Scheiner-Bobis, G., and Schreiber, S. (1999) *Biochemistry* **38**, 9198–9208
534. Gatto, C., Wang, A. X., and Kaplan, J. H. (1998) *J. Biol. Chem.* **273**, 10578–10585
535. Rossier, B. C., Geering, K., and Kraehenbuhl, J. P. (1987) *Trends Biochem. Sci.* **12**, 483–487
536. Minor, N. T., Sha, Q., Nichols, C. G., and Mercer, R. W. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6521–6525
537. Lutsenko, S., and Kaplan, J. H. (1995) *Biochemistry* **34**, 15607–15613
538. Nelson, N., and Taiz, L. (1989) *Trends Biochem. Sci.* **14**, 113–116
539. Solioz, M., and Vulpe, C. (1996) *Trends Biochem. Sci.* **21**, 237–241

539a. Ueno, S., Kaieda, N., Koyama, N. (2000) *J. Biol. Chem.* **275**, 14537–14540

540. Carafoli, E. (1994) *FASEB J.* **8**, 993–1002
541. MacLennan, D. H., Rice, W. J., and Green, N. M. (1997) *J. Biol. Chem.* **272**, 28815–28818
542. Myung, J., and Jencks, W. P. (1995) *Biochemistry* **34**, 3077–3083
543. Dode, L., De Greef, C., Mountian, I., Attard, M., Town, M. M., Casteels, R., and Wuytack, F. (1998) *J. Biol. Chem.* **273**, 13982–13994

543a. Menguy, T., Corre, F., Bouneau, L., Deschamps, S., Vuust Moller, J., Champeil, P., le Maire, M., and Falson, P. (1998) *J. Biol. Chem.* **273**, 20134–20143

544. Melle-Milovanovic, D., Milovanovic, M., Nagpal, S., Sachs, G., and Shin, J. M. (1998) *J. Biol. Chem.* **273**, 11075–11081
545. Rulli, S. J., Horiba, M. N., Skripnikova, E., and Rabon, E. C. (1999) *J. Biol. Chem.* **274**, 15245–15250
546. Asano, S., Inoie, M., and Takeguchi, N. (1987) *J. Biol. Chem.* **262**, 13263–13268
547. Seto-Young, D., Hall, M. J., Na, S., Haber, J. E., and Perlin, D. S. (1996) *J. Biol. Chem.* **271**, 581–587
548. Auer, M., Scarborough, G. A., and Kühlbrandt, W. (1998) *Nature (London)* **392**, 840–843
549. Orlowski, J., and Grinstein, S. (1997) *J. Biol. Chem.* **272**, 22373–22376
549. Orlowski, J. and Grinstein, S. (1997) *J. Biol. Chem.* **272**, 22373–22376

549a. Wakabayashi, S., Pang, T., Su, X., and Shigekawa, M. (2000) *J. Biol. Chem.* **275**, 7942–7949

549b. Williams, K. A. (2000) *Nature (London)* **403**, 112–115

550. Kasho, V. N., Stengelin, M., Smirnova, I. N., and Faller, L. D. (1997) *Biochemistry* **36**, 8045–8052
551. Campos, M., and Beaugé, L. (1997) *Biochemistry* **36**, 14228–14237
552. Vilsen, B. (1997) *Biochemistry* **36**, 13312–13324

552a. Boxenbaum, N., Daly, S. E., Javaid, Z. Z., Lane, L. K., and Blostein, R. (1998) *J. Biol. Chem.* **273**, 23086–23092

553. Zhang, P., Toyoshima, C., Yonekura, K., Green, N. M., and Stokes, D. L. (1998) *Nature (London)* **392**, 835–839

553a. Toyoshima, C., Nakasako, M., Nomura, H., and Ogawa, H. (2000) *Nature (London)* **405**, 647–655

## References


553b. MacLennan, D. H., and Green, N. M. (2000) *Nature (London)* **405**, 633–634
553c. Zhang, Z., Lewis, D., Strock, C., and Inesi, G. (2000) *Biochemistry* **39**, 8758–8767
554. Crider, B. P., Andersen, P., White, A. E., Zhou, Z., Li, X., Mattsson, J. P., Lundberg, L., Keeling, D. J., Xie, X.-S., Stone, D. K., and Peng, S.-B. (1997) *J. Biol. Chem.* **272**, 10721–10728
555. Nanda, A., Brumell, J. H., Nordström, T., Kjeldsen, L., Sengelov, H., Borregaard, N., Rotstein, O. D., and Grinstein, S. (1996) *J. Biol. Chem.* **271**, 15963–15970
555a. Kelly, R. B. (1999) *Trends Biochem. Sci.* **24**, M29–M33
556. Desjardins, M., Celis, J. E., van Meer, G., Dieplinger, H., Jahraus, A., Griffiths, G., and Huber, L. A. (1994) *J. Biol. Chem.* **269**, 32194–32200
557. Peng, S.-B., Crider, B. P., Tsai, S. J., Xie, X.-S., and Stone, D. K. (1996) *J. Biol. Chem.* **271**, 3324–3327
558. Pearse, B. M. F., and Crowther, R. A. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 49–68
559. Schmid, S. L., and Damke, H. (1995) *FASEB J.* **9**, 1445–1453
559a. Marsh, M., and McMahon, H. T. (1999) *Science* **285**, 215–220
559b. Heuser, J. (1981) *Trends Biochem. Sci.* **6**, 64–68
560. Pearse, B. M. F., and Bretscher, M. S. (1981) *Ann. Rev. Biochem.* **50**, 85–101
561. Pishvaee, B., Munn, A., and Payne, G. (1997) *EMBO J.* **16**, 2227–2239
562. Rothman, J. E. (1996) *Protein Sci.* **5**, 185–194
563. Brodsky, F. M., Hill, B. L., Acton, S. L., Näthke, I., Wong, D. H., Ponnambalam, S., and Parham, P. (1991) *Trends Biochem. Sci.* **16**, 208–213
564. Lemmon, S. K., and Jones, E. W. (1987) *Science* **238**, 504–509
564a. Anderson, R. G. W. (1998) *Ann. Rev. Biochem.* **67**, 199–225
564b. Shin, J.-S., Gao, Z., and Abraham, S. N. (2000) *Science* **289**, 785–788
564c. Mulvey, M. A., and Hultgren, S. J. (2000) *Science* **289**, 732–733
565. Lowe, M., and Kreis, T. E. (1995) *J. Biol. Chem.* **270**, 31364–31371
566. Schekman, R., and Orci, L. (1996) *Science* **271**, 1526–1533
566a. Eugster, A., Frigerio, G., Dale, M., and Duden, R. (2000) *EMBO J.* **19**, 3905–3917
567. Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R., and Klug, A. (1978) *Nature (London)* **276**, 362–368
567a. Rapoport, I., Chen, Y. C., Cupers, P., Shoelson, S. E., and Kirchhausen, T. (1998) *EMBO J.* **17**, 2148–2155
567b. Huang, K. M., D'Hondt, K., Riezman, H., and Lemmon, S. K. (1999) *EMBO J.* **18**, 3897–3908
567c. Owen, D. J., Vallis, Y., Pearse, B. M. F., McMahon, H. T., and Evans, P. R. (2000) *EMBO J.* **19**, 4216–4227
567d. Wendland, B., Steece, K. E., and Emr, S. D. (1999) *EMBO J.* **18**, 4383–4393
567e. Scales, S. J., and Scheller, R. H. (1999) *Nature (London)* **401**, 123–124
568. Ungewickell, E., Ungewickell, H., Holstein, S. E. H., Lindner, R., Prasad, K., Barouch, W., Martin, B., Greene, L. E., and Eisenberg, E. (1995) *Nature (London)* **378**, 632–635
569. Mellman, I., Fuchs, R., and Helenius, A. (1986) *Ann. Rev. Biochem.* **55**, 663–700
570. Brown, M. S., and Goldstein, J. L. (1974) *Sci. Am.* **251**(Nov), 58–66
571. Brown, M. S., and Goldstein, J. L. (1986) *Science* **232**, 34–47
572. Holstein, T. W., Benoit, M., Herder, G., Wanner, G., David, C. N., and Gaub, H. E. (1994) *Science* **265**, 402–404
573. Ferro-Novick, S., and Jahn, R. (1994) *Nature (London)* **370**, 191–193
574. Mayorga, L. S., Berón, W., Sarrouf, M. N., Colombo, M. I., Creutz, C., and Stahl, P. D. (1994) *J. Biol. Chem.* **269**, 30927–30934
575. Burgoyne, R. D. (1988) *Nature (London)* **331**, 20
576. Wilson, D. W., Wilcox, C. A., Flynn, G. C., Chen, E., Kuang, W.-J., Henzel, W. J., Block, M. R., Ullrich, A., and Rothman, J. E. (1989) *Nature (London)* **339**, 355–359
577. Colombo, M. I., Taddese, M., Whiteheart, S. W., and Stahl, P. D. (1996) *J. Biol. Chem.* **271**, 18810–18816
578. Goda, Y. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 769–772
579. Lugtenberg, B. (1981) *Trends Biochem. Sci.* **6**, 262–266
580. Graham, L. L., Beveridge, T. J., and Nanninga, N. (1991) *Trends Biochem. Sci.* **16**, 328–329
581. Oliver, D. B. (1987) in Escherichia coli *and* Salmonella typhi-murium (Neidhardt, F. C., ed), pp. 56–70, Am. Soc. for Micro-biology, Washington, DC
582. Glauner, B., Höltje, J.-V., and Schwarz, U. (1988) *J. Biol. Chem.* **263**, 10088–10095
583. Park, J. T. (1987) in Escherichia coli *and* Salmonella typhimurium (Neidhardt, F. C., ed), pp. 23–32, Am. Soc. for Microbiology, Washington, DC
584. Scherrer, R. (1984) *Trends Biochem. Sci.* **9**, 242–245
585. Nikaido, H., and Vaara, M. (1987) in Escherichia coli *and* Salmonella typhimurium (Neidhardt, F. C., ed), pp. 7–22, Am. Soc. for Microbiology, Washington, DC
586. Hancock, R. E. W., Karunaratne, D. N., and Bernegger-Egli, C. (1994) in *Bacterial Cell Wall (New Comprehensive Biochemistry)*, Vol. 27 (Ghuysen, J.-M., and Hakenbeck, R., eds), pp. 263–279, Elsevier, Amsterdam
587. Raetz, C. R. H., and Dowhan, W. (1990) *J. Biol. Chem.* **265**, 1235–1238
588. Osborn, M. J. (1969) *Ann. Rev. Biochem.* **38**, 501–538
589. Choi, D.-S., Yamada, H., Mizuno, T., and Mizushima, S. (1986) *J. Biol. Chem.* **261**, 8953–8957
590. Diedrich, D. L., and Schnaitman, C. A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3708–3712
591. Raetz, C. R. H., Ulevitch, R. J., Wright, S. D., Sibley, C. H., Ding, A., and Nathan, C. F. (1991) *FASEB J.* **5**, 2652–2660
592. Hollingsworth, R. I., and Carlson, R. W. (1989) *J. Biol. Chem.* **264**, 9300–9303
593. Ogawa, T. (1994) *Eur. J. Biochem.* **219**, 737–742
594. Rietschel, E. T., Kirikae, T., Schade, F. U., Mamat, U., Schmidt, G., Loppnow, H., Ulmer, A. J., Zähringer, U., Seydel, U., Di Padova, F., Schreier, M., and Brade, H. (1994) *FASEB J.* **8**, 217–225
595. Masoud, H., Moxon, E. R., Martin, A., Krajcarski, D., and Richards, J. C. (1997) *Biochemistry* **36**, 2091–2103
596. Aspinall, G. O., and Monteiro, M. A. (1996) *Biochemistry* **35**, 2498–2504
597. Rietschel, E. T., and Brade, H. (1992) *Sci. Am.* **267**(Aug), 54–61
598. Costerton, J. W., Irwin, R. T., and Cheng, K. J. (1981) *Crit. Rev. Microbiol* **8**, 303–338
599. Fischer, W. (1994) in *Bacterial Cell Wall (New Comprehensive Biochemistry)*, Vol. 27 (Ghuysen, J.-M., and Hakenbeck, R., eds), pp. 199–215, Elsevier, Amsterdam
600. Ganfield, M.-C. W., and Pieringer, R. A. (1980) *J. Biol. Chem.* **255**, 5164–5169
601. Johnson, S. D., Lacher, K. P., and Anderson, J. S. (1981) *Biochemistry* **20**, 4781–4785
602. Ivatt, R. J., and Gilvarg, C. (1979) *J. Biol. Chem.* **254**, 2759–2765
603. Venisse, A., Rivière, M., Vercauteren, J., and Puzo, G. (1995) *J. Biol. Chem.* **270**, 15012–15021
604. Brennan, P. J., and Nikaido, H. (1995) *Ann. Rev. Biochem.* **64**, 29–63
605. Liu, J., Barry, C. E., III, Besra, G. S., and Nikaido, H. (1996) *J. Biol. Chem.* **271**, 29545–29551
606. Mikusová, K., Mikus, M., Besra, G. S., Hancock, I., and Brennan, P. J. (1996) *J. Biol. Chem.* **271**, 7820–7828
607. Lopez-Marin, L. M., Quesada, D., Lakhdar-Ghazal, F., Tocanne, J.-F., and Lanéelle, G. (1994) *Biochemistry* **33**, 7056–7061
608. Sprott, G. D., Ekiel, I., and Dicaire, C. (1990) *J. Biol. Chem.* **265**, 13735–13740
609. Peters, J., Nitsch, M., Kühlmorgen, B., Golbik, R., Lupas, A., Kellermann, J., Engelhardt, H., Pfander, J.-P., Müller, S., Goldie, K., Engel, A., Stetter, K.-O., and Baumeister, W. (1995) *J. Mol. Biol.* **245**, 385–401
610. Thomas, S. R., and Trust, T. J. (1995) *J. Mol. Biol.* **245**, 568–581
611. Peters, J., Baumeister, W., and Lupas, A. (1996) *J. Mol. Biol.* **257**, 1031–1041
612. Fischetti, V. A. (1991) *Sci. Am.* **264**(Jun), 58–65
613. Goward, C. R., Scawen, M. D., Murphy, J. P., and Atkinson, T. (1993) *Trends Biochem. Sci.* **18**, 136–139
614. Schneewind, O., Fowler, A., and Faull, K. F. (1995) *Science* **268**, 103–106
615. Knörle, R., and Hübner, W. (1995) *Biochemistry* **34**, 10970–10975
616. Shaper, N. L., Hollis, G. F., Douglas, J. G., Kirsch, I. R., and Shaper, J. H. (1988) *J. Biol. Chem.* **263**, 10420–10428
617. Lin, C. Q., and Bissell, M. J. (1993) *FASEB J.* **7**, 737–743
618. Har-El, R., and Tanzer, M. L. (1993) *FASEB J.* **7**, 1115–1123
619. Mayne, R., and Burgeson, R. E., eds. (1987) *Structure and Function of Collagen Types*, Academic Press, New York, Orlando, San Diego
620. Nimni, M. E., ed. (1988) *Collagen: Biochemistry, Biomechanics, Biotechnology*, CRC Press, Boca Raton, Florida
621. Ramachandran, G. N. (1988) *Intl. J. Peptide Prot. Res.* **31**, 1–16
622. Martin, G. R., Timpl, R., Müller, P. K., and Kühn, K. (1985) *Trends Biochem. Sci.* **10**, 285–287
623. Prockop, D. J., and Kivirikko, K. I. (1995) *Ann. Rev. Biochem.* **64**, 403–434
624. van der Rest, M., and Garrone, R. (1991) *FASEB J.* **5**, 2814–2823
625. Byers, P. H. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4029–4077, McGraw-Hill, New York
626. Brodsky, B., and Shah, N. K. (1995) *FASEB J.* **9**, 1537–1546
627. Koivu, J., and Myllylä, R. (1987) *J. Biol. Chem.* **262**, 6159–6164
628. Prockop, D. J. (1990) *J. Biol. Chem.* **265**, 15349–15352
629. Annunen, P., Autio-Harmainen, H., and Kivirikko, K. I. (1998) *J. Biol. Chem.* **273**, 5989–5992
630. Reiser, K., McCormick, R. J., and Rucker, R. B. (1992) *FASEB J.* **6**, 2439–2449
631. Li, S.-W., Sieron, A. L., Fertala, A., Hojima, Y., Arnold, W. V., and Prockop, D. J. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5127–5130
632. Meek, K. M., Chapman, J. A., and Hardcastle, R. A. (1979) *J. Biol. Chem.* **254**, 10710–10714
633. Chan, V. C., Ramshaw, J. A. M., Kirkpatrick, A., Beck, K., and Brodsky, B. (1997) *J. Biol. Chem.* **272**, 31441–31446

## References


634. Wess, T. J., Hammersley, A. P., Wess, L., and Miller, A. (1998) *J. Mol. Biol.* **275**, 255–267
634a. Kramer, R. Z., Venugopal, M. G., Bella, J., Mayville, P., Brodsky, B., and Berman, H. M. (2000) *J. Mol. Biol.* **301**, 1191–1205
635. Eyre, D. R., Paz, M. A., and Gallop, P. M. (1984) *Ann. Rev. Biochem.* **53**, 717–748
636. Miller, A. (1982) *Trends Biochem. Sci.* **7**, 13–18
637. Eyre, D. R. (1980) *Science* **207**, 1315–1322
638. Wu, J.-J., and Eyre, D. R. (1984) *Biochemistry* **23**, 1850–1857
639. Hanson, D. A., and Eyre, D. R. (1996) *J. Biol. Chem.* **271**, 26508–26516
640. Kiss, I., Deák, F., Holloway, R. G., Jr., Delius, H., Mebust, K. A., Frinberger, E., Argraves, W. W., Tsonis, P. A., Winterbottom, N., and Goetink, P. F. (1989) *J. Biol. Chem.* **264**, 8126–8134
641. Tanaka, S., Avigad, G., Eikenberry, E. F., and Brodsky, B. (1988) *J. Biol. Chem.* **263**, 17650–17657
642. Kleman, J.-P., Aeschlimann, D., Paulsson, M., and van der Rest, M. (1995) *Biochemistry* **34**, 13768–13775
643. Mazzorana, M., Gruffat, H., Sergeant, A., and van der Rest, M. (1993) *J. Biol. Chem.* **268**, 3029–3032
644. Martin, G. R., Timpl, R., Miller, P. K., and Kuhn, K. (1985) *Trends Biochem. Sci.* **10**, 285–287
645. Gunwar, S., Ballester, F., Noelken, M. E., Sado, Y., Ninomiya, Y., and Hudson, B. G. (1998) *J. Biol. Chem.* **273**, 8767–8775
646. van der Rest, M., and Mayne, R. (1988) *J. Biol. Chem.* **263**, 1615–1618
647. de Crombrugghe, B., and Pastan, I. (1982) *Trends Biochem. Sci.* **7**, 11–13
648. Myers, J. C., Yang, H., D'Ippolito, J. A., Presente, A., Miller, M. K., and Dion, A. S. (1994) *J. Biol. Chem.* **269**, 18549–18557
649. Lui, V. C. H., Ng, L. J., Sat, E. W. Y., Nicholls, J., and Cheah, K. S. E. (1996) *J. Biol. Chem.* **271**, 16945–16951
650. de Wet, W., Bernard, M., Benson-Chanda, V., Chu, M.-L., Dickson, L., Weil, D., and Ramirez, F. (1987) *J. Biol. Chem.* **262**, 16032–16036
651. Chu, M.-L., de Wet, W., Bernhard, M., and Ramirez, F. (1985) *J. Biol. Chem.* **260**, 2315–2370
652. Kramer, J. M. (1994) *FASEB J.* **8**, 329–336
653. Engel, J. (1997) *Science* **277**, 1785–1786
654. Rosenbloom, J., Abrams, W. R., and Mecham, R. (1993) *FASEB J.* **7**, 1208–1218
655. Hukins, D. W. L. (1984) *Connective Tissue Matrix*, MacMillan, New York
656. Fransson, L.-Å. (1987) *Trends Biochem. Sci.* **12**, 406–411
657. Burg, M. A., Tillet, E., Timpl, R., and Stallcup, W. B. (1996) *J. Biol. Chem.* **271**, 26110–26116
658. Caplan, A. I. (1984) *Sci. Am.* **251**(Oct), 84–94
659. Heinegård, D., and Oldberg, Å. (1989) *FASEB J.* **3**, 2042–2051
660. Cheng, F., Heinegård, D., Fransson, L.-Å., Bayliss, M., Bielicki, J., Hopwood, J., and Yoshida, K. (1996) *J. Biol. Chem.* **271**, 28572–28580
661. Hauser, N., and Paulsson, M. (1994) *J. Biol. Chem.* **269**, 25747–25753
662. Beck, K., Gambee, J. E., Bohan, C. A., and Bächinger, H. P. (1996) *J. Mol. Biol.* **256**, 909–923
663. Yurchenco, P. D., and Schittny, J. C. (1990) *FASEB J.* **4**, 1577–1590
663a. Yurchenco, P. D., Birk, D. E., and Mecham, R. P., eds. (1994) *Extracellular Matrix Assembly and Structure*, Academic Press, San Diego, California
664. Utani, A., Nomizu, M., Timpl, R., Roller, P. P., and Yamada, Y. (1994) *J. Biol. Chem.* **269**, 19167–19175
665. Pöschl, E., Mayer, U., Stetefeld, J., Baumgartner, R., Holak, T. A., Huber, R., and Timpl, R. (1996) *EMBO J.* **15**, 5154–5159
666. Durkin, M. E., Gautam, M., Loechel, F., Sanes, J. R., Merlie, J. P., Albrechtsen, R., and Wewer, U. M. (1996) *J. Biol. Chem.* **271**, 13407–13416
667. Beckmann, G., Hanke, J., Bork, P., and Reich, J. G. (1998) *J. Mol. Biol.* **275**, 725–730
668. Colognato-Pyke, H., O'Rear, J. J., Yamada, Y., Carbonetto, S., Cheng, Y.-S., and Yurchenco, P. D. (1995) *J. Biol. Chem.* **270**, 9398–9406
669. Kramer, J. M., Cox, G. N., and Hirsch, D. (1985) *J. Biol. Chem.* **260**, 1945–1951
670. Beck, K., Hunter, I., and Engel, J. (1990) *FASEB J.* **4**, 148–160
671. Sasaki, T., Kostka, G., Göhring, W., Wiedemann, H., Mann, K., Chu, M.-L., and Timpl, R. (1995) *J. Mol. Biol.* **245**, 241–250
672. Aeschlimann, D., and Paulsson, M. (1991) *J. Biol. Chem.* **266**, 15308–15317
673. Reinhardt, D. P., Ono, R. N., and Sakai, L. Y. (1997) *J. Biol. Chem.* **272**, 1231–1236
673a. Kettle, S., Yuan, X., Grundy, G., Knott, V., Downing, A. K., and Handford, P. A. (1999) *J. Mol. Biol.* **285**, 1277–1287
673b. Ritty, T. M., Broekelmann, T., Tisdale, C., Milewicz, D. M., and Mecham, R. P. (1999) *J. Biol. Chem.* **274**, 8933–8940
674. Reinhardt, D. P., Keene, D. R., Corson, G. M., Pöschl, E., Bächinger, H. P., Gambee, J. E., and Sakai, L. Y. (1996) *J. Mol. Biol.* **258**, 104–116
675. McKusick, V. A. (1991) *Nature (London)* **352**, 279–281
676. Francke, U., and Furthmayr, H. (1994) *N. Engl. J. Med.* **330**, 1384–1385
677. Mann, K., Mechling, D. E., Bächinger, H. P., Eckerskorn, C., Gaill, F., and Timpl, R. (1996) *J. Mol. Biol.* **261**, 255–266
678. Durkin, M. E., Carlin, B. E., Vergnes, J., Bartos, B., and Merlie, J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1570–1574
679. Richards, A. G. (1978) in *Biochemistry of Insects* (Rockstein, M., ed), Academic Press, New York (p. 205)
680. Schaefer, J., Kramer, K. J., Garbow, J. R., Jacob, G. S., Stejskal, E. O., Hopkins, T. L., and Speirs, R. D. (1987) *Science* **235**, 1200–1204
681. Simkiss, K., and Wilbur, K. M. (1989) *Biomineralization*, Academic Press, San Diego, California
682. DeOliveira, D. B., and Laursen, R. A. (1997) *J. Am. Chem. Soc.* **119**, 10627–10631
683. Weiner, S., and Addadi, L. (1991) *Trends Biochem. Sci.* **16**, 252–256
684. Albeck, S., Aizenberg, J., Addadi, L., and Weiner, S. (1993) *J. Am. Chem. Soc.* **115**, 11691–11697
685. Bonucci, E., ed. (1992) *Calcification in Biological Systems*, CRC Press, Boca Raton, Florida
686. Mann, S., Archibald, D. D., Didymus, J. M., Douglas, T., Heywood, B. R., Meldrum, F. C., and Reeves, N. J. (1993) *Science* **261**, 1286–1292
687. Aizenberg, J., Hanson, J., Ilan, M., Leiserowitz, L., Koetzle, T. F., Addadi, L., and Weiner, S. (1995) *FASEB J.* **9**, 262–268
688. Otsubo, K., Katz, E. P., Mechanic, G. L., and Yamauchi, M. (1992) *Biochemistry* **31**, 396–402
689. Fledelius, C., Johnsen, A. H., Cloos, P. A. C., Bonde, M., and Qvist, P. (1997) *J. Biol. Chem.* **272**, 9755–9763
689a. Ducy, P., Schinke, T., and Karsenty, G. (2000) *Science* **289**, 1501–1504
689b. Teitelbaum, S. L. (2000) *Science* **289**, 1504–1508
690. Weiner, S., and Traub, W. (1992) *FASEB J.* **6**, 879–885
691. McLean, F. M., Keller, P. J., Genge, B. R., Walters, S. A., and Wuthier, R. E. (1987) *J. Biol. Chem.* **262**, 10481–10488
692. Komori, T., Yagi, H., Nomura, S., Yamaguchi, A., Sasaki, K., Deguchi, K., Shimizu, Y., Bronson, R. T., Gao, Y.-H., Inada, M., Sato, M., Okamoto, R., Kitamura, Y., Yoshiki, S., and Kishimoto, T. (1997) *Cell* **89**, 755–764
693. Ducy, P., Zhang, R., Geoffroy, V., Ridall, A. L., and Karsenty, G. (1997) *Cell* **89**, 747–754
694. Dickman, S. (1997) *Science* **276**, 1502
695. Rodan, G. A., and Harada, S.-i. (1997) *Cell* **89**, 677–680
696. Mundlos, S., Otto, F., Mundlos, C., Mulliken, J. B., Aylsworth, A. S., Albright, S., Lindhout, D., Cole, W. G., Henn, W., Knoll, J. H. M., Owen, M. J., Mertelsmann, R., Zabel, B. U., and Olsen, B. R. (1997) *Cell* **89**, 773–779
697. Price, P. A., and Williamson, M. K. (1985) *J. Biol. Chem.* **260**, 14971–14975
698. Nishimoto, S. K., Araki, N., Robinson, F. D., and Waite, J. H. (1992) *J. Biol. Chem.* **267**, 11600–11605
699. Price, P. A., Rice, J. S., and Williamson, M. K. (1994) *Protein Sci.* **3**, 822–830
700. Hohenester, E., Maurer, P., and Timpl, R. (1997) *EMBO J.* **16**, 3778–3786
701. Kelm, R. J., Jr., Swords, N. A., Orfeo, T., and Mann, K. G. (1994) *J. Biol. Chem.* **269**, 30147–30153
702. Fisher, L. W., Robey, P. G., Tuross, N., Otsuka, A. S., Tepen, D. A., Esch, F. S., Shimasaki, S., and Termine, J. D. (1987) *J. Biol. Chem.* **262**, 13457–13463
703. Denhardt, D. T., and Guo, X. (1993) *FASEB J.* **7**, 1475–1482
704. Shanmugam, V., Chackalaparampil, I., Kundu, G. C., Mukherjee, A. B., and Mukherjee, B. B. (1997) *Biochemistry* **36**, 5729–5738
704a. Rodan, G. A., and Martin, T. J. (2000) *Science* **289**, 1508–1514
704b. Service, R. F. (2000) *Science* **289**, 1498–1500
705. Zipkin, I., ed. (1973) *Biological Mineralization*, Wiley, New York
706. Schlesinger, P. H., Blair, H. C., Teitelbaum, S. L., and Edwards, J. C. (1997) *J. Biol. Chem.* **272**, 18636–18643
707. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994) *Molecular Biology of the Cell*, 3rd ed., Garland, New York (p. 1182–1186)
708. Riggs, B. L., and Melton, L. J. (1992) *N. Engl. J. Med.* **327**, 620–627
709. Sly, W. S., and Hu, P. Y. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., Vol. 3 (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds), pp. 4113–4124, McGraw-Hill, New York
710. Roush, W. (1996) *Science* **273**, 1170
711. Delmas, P. D., and Meunier, P. J. (1997) *N. Engl. J. Med.* **336**, 558–566
711a. Ho, A. M., Johnson, M. D., and Kingsley, D. M. (2000) *Science* **289**, 265–270
711b. Hagmann, M. (2000) *Science* **289**, 225–226
712. Hohling, H. J., Barckhaus, R. H., and Kreftin, E. R. (1980) *Trends Biochem. Sci.* **5**, 8–11
713. Outka, D. E., and Williams, D. C. (1971) *Journal of Protozoology* **18**, 285–297
714. Inoue, S., and Okazaki, K. (1977) *Sci. Am.* **236**(Apr), 83–92
715. Volcani, B. E. (1981) in *Silicon and Silicious Structures in Biological Systems* (Simpson, T. L., and Volcani, B. E., eds), Springer-Verlag, New York (p. 157)
716. Mann, S. (1988) *Nature (London)* **332**, 119–124
717. Calvert, P. (1988) *Nature (London)* **334**, 651–652
718. Watanabe, N., and Wilbur, K. M., eds. (1976) *The Mechanisms of Mineralization in Invertebrates and Plants*, Univ. N. Carolina Press, Columbia, South Carolina

## References


719. Lemann, J., Jr. (1993) *N. Engl. J. Med.* **328**, 880–881
720. Nakagawa, Y., Abram, V., Kézdy, F. J., Kaiser, E. T., and Coe, F. L. (1983) *J. Biol. Chem.* **258**, 12594–12600
721. Bertrand, J. A., Pignol, D., Bernard, J.-P., Verdier, J.-M., Dagorn, J.-C., and Fontecilla-Camps, J. C. (1996) *EMBO J.* **15**, 2678–2684
722. Bimmler, D., Graf, R., Scheele, G. A., and Frick, T. W. (1997) *J. Biol. Chem.* **272**, 3073–3082
723. De Reggi, M., Gharib, B., Patard, L., and Stoven, V. (1998) *J. Biol. Chem.* **273**, 4967–4971
724. Cheung, H. S., Kurup, I. V., Sallis, J. D., and Ryan, L. M. (1996) *J. Biol. Chem.* **271**, 28082–28085
725. Rothschild, B. M., and Thillaud, P. L. (1991) *Nature (London)* **349**, 288
726. Cabib, E., Roberts, R., and Bowers, B. (1982) *Ann. Rev. Biochem.* **51**, 763–793
727. Ballou, C. E., and Raschke, W. C. (1974) *Science* **184**, 127–134
728. Frevert, J., and Ballou, C. E. (1985) *Biochemistry* **24**, 753–759
729. Celerin, M., Ray, J. M., Schisler, N. J., Day, A. W., Stetler-Stevenson, W. G., and Laudenbach, D. E. (1996) *EMBO J.* **15**, 4445–4453


## Study Questions

1. Compare the chemical makeup of the extracellar "coat" or "matrix" materials secreted by the following cells: bacteria, fibroblasts, osteoblasts, plant cells, fungi.

2. The iodine number of a compound is defined as the number of grams of $I_2$ absorbed (through addition to C=C bonds to give a diiodo derivative) per 100 g of fat. NOTE: Iodine monochloride or iodine monobromide are the usual halogenating reagents but the iodine number is expressed in terms of grams of $I_2$. The saponification number is the number of milligrams of KOH needed to completely saponify (hydrolyze and neutralize the resulting fatty acids) 1 g of fat. A pure triglyceride has a saponification number of 198 and an iodine number of 59.7.

   a) What is the relative molecular mass?
   b) What is the average chain length of the fatty acids?
   c) What is the number of double bonds in the molecule?

3. Spermaceti (a wax from the head of the sperm whale) resembles high molecular mass hydrocarbons in physical properties and inertness toward $Br_2$ / $CHCl_3$ and $KMNO_4$; on qualitative analysis, it gives positive tests only for carbon and hydrogen. However, its IR spectrum shows the presence of an ester linkage, and quantitative analysis gives the empirical formula $C_{16}H_{32}O$. A solution of the wax in alcoholic KOH is refluxed for a long time. Titration of an aliquot shows that one equivalent of base is consumed for every 475 g of wax. Water and ether are added to the cooled reflux mixture, and the aqueous and etheral layers are separated. Acidification of the aqueous layer yields a solid A with a neutralization equivalent of 260 $\pm$ 5. Evaporation of the ether layer gives solid B, which could not be titrated. Reduction of either spermaceti or A by lithium aluminum hydride gave B as the only product.

   What is the likely structure of spermaceti?

4. Stearic acid (1.16 g) was dissolved in 100 ml of ethanol. A 10 µl portion of the resulting solution was pipetted onto a clean surface of a dilute HCl solution (in a shallow tray) where it spread to form a monolayer of stearic acid. The layer was compressed (by moving a Teflon barrier across the tray) until the surface pressure $\pi$ started to rise sharply and reached ~20 dyn / cm. Note that $\pi = \gamma_o - \gamma$ where $\gamma$ is the measured surface tension with the film present and $\gamma_o$ is the higher surface tension of water alone. The compressed film occupied a 20 x 24 cm area. Calculate the cross-sectional area of an alkyl chain in stearic acid. [See J. B. Davenport, *in Biochemistry and Methodology of Lipids* (A. R. Johnson and J. B. Davenport, eds.), pp. 47-83. Wiley-Intersience, New York, 1971; and M. C. Phillips, *in Progress in Surface and Membrane Science* (J. F. Danielli, D. M. Rosenberg, and D. A. Cadenhead, eds.) Vol. 5, pp. 139-221. Academic Press, New York, 1972.]

5. In 1925, E. Gorter and F. Grendel (*J. Exp. Med.* **41**, 439) reported measurements in which they extracted lipid from red blood cell membranes with acetone, spread the lipids as a monolayer, and measured the area of the compressed monolayer. They then estimated the surface area of an erythrocyte and calculated that the ratio of the lipids (as a monolayer) to the surface area of the red blood cell was 1.9–2.0. More modern experiments gave the following: each erythrocyte membrane contains 4.5 x $10^{-16}$ mol of phospholipid and 3.1 x $10^{-16}$ mol of cholesterol.

   a) If the cross-sectional areas of phospholipid and cholesterol molecules in a membrane are taken as 0.70 and 0.38 $nm^2$, respectively, what surface area would be occupied in a monolayer?
   b) If the measured surface area of an erythrocyte is 167 $\mu m^2$, what is the ratio of the area calculated in (a) to the area of the cell surface?
   c) How might you explain the difference between this answer and that of Gorter and Grendel? See E. D. Korn (1966) *Science* **153**, 1491–1498.

## Study Questions

6. The following experimental observations are related to biological membrane structure and function. Discuss the implications of each observation with respect to membrane structure.

   a) Many macrocyclic antibiotics (nonactin, valinomycin, and others) form 1: 1 complexes with alkali metal cations in a highly selective manner. The complexes are readily soluble in nonpolar organic solvents. These antibiotics increase the electrical conductance and permeability to alkali metal cations of synthetic phospholipid membranes. Valinomycin increases the electrical conductance of thylakoid membranes of chloroplasts in the presence of $K^+$ but not in the presence of $Na^+$; it also uncouples oxidative phoshorylation in mitochondria. (See Chapter 18.)
   b) Treatment of intact chloroplasts with a galactolipase releases glactose from galactosyl diglycerides. Treatment of red blood cell ghosts with phopholipase C releases about 75% of the lipid P in water-soluble form. In neither case is the structural integrity of the membrane destroyed.
   c) Using sphingomyelin as a hapten, antibodies specific for this lipid can be produced. When red blood cell ghosts are exposed to these antibodies, it can be shown that the antibodies react, but only on one side of the membrane.

7. Describe the structure of biological membranes and the characteristic functions of lipid-, protein-, and carbohydrate-containing components. Describe the differences between inner and outer membrane surfaces.

8. Compare the distribution of triglycerides, phosphatidylcholine, phosphatidylethanolamine, sphingomyelins, glycolipids, and cholesterol within cells. Consider differences between the two sides of membranes.

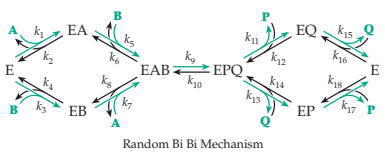
9. Consider the chemistry underlying the labeling of cell surfaces with each of the following:
   a) Lactoperoxidase,
   b) galactose oxidase,
   c) formylmethionylsulfone methyl phosphate,
   d) the diazonium salt of diiodosulfanilic acid,
   e) fluorescent antibodies,
   f) antibodies conjugated with ferritin.

   Write the equations for the chemical reactions involved. State what surface groups will be labeled by each reagent. List special advantages of each of these reagents.

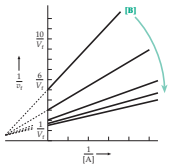
10. Which would be the more effective detergent in the pH range 2 to 3, sodium lauryl sulfonate or sodium laurate? Why?

11. Suppose that a cell contains 10 mM $Na^+$ and 100 mM $K^+$ and that it is bathed in extracellular fluid containing 100 mM $Na^+$ and 5 mM $K^+$. How much energy will be required to transport three equivalents of $Na^+$ out and two equivalents of $K^+$ in? Compare this with $\Delta G'$ for hydrolysis of ATP at pH 7. Assume that the membrane is permeable to $Cl^-$.

12. An *E. coli* cell is said to contain about $10^5$ molecules of an envelope protein of MW = 36,500. If the latter is spherical and the spheres are closely packed in a hexogonal lattice, how much of the surface area of the bacterium would be covered? What would the diameter of the protein be? What spacing would be required if $10^5$ molecules covered the surface completely? Suggest a shape for the protein molecule that is consistent with the requirement.



Random Bi Bi Mechanism




Bimolecular reactions of two molecules, A and B, to give two products, P and Q, are catalyzed by many enzymes. For some enzymes the substrates A and B bind into the active site in an ordered sequence while for others, binding may be in a random order. The scheme shown here is described as random Bi Bi in a classification introduced by Cleland. Eighteen rate constants, some second order and some first order, describe the reversible system. Determination of these kinetic parameters is often accomplished using a series of double reciprocal plots (Lineweaver-Burk plots), such as those at the right.

# Contents

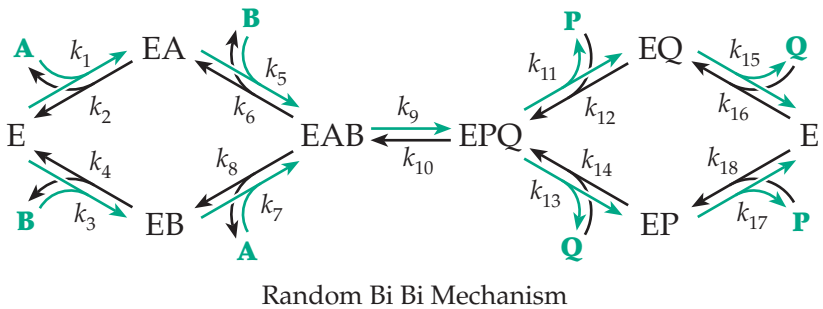

455 **A. Information from Kinetics**
455 1. Measuring the Speed of an Enzymatic Reaction
457 *First-order reactions*
457 *Turnover numbers and units of activity*
458 *Second-order reactions*
458 *Reversible chemical reactions*
458 2. Formation and Reaction of Enzyme–Substrate Complexes
460 *Linear forms for rate equations*
460 *Nonlinear equations and integrated rate equations*
461 *Kinetics with high enzyme concentrations*
461 3. Diffusion and the Rate of Encounter of an Enzyme with Substrate
462 *The "cage effect" and rotation of molecules*
463 *The rate of substrate binding*
463 *The displacement of bound ligands by substrate*
464 4. Reversible Enzymatic Reactions
464 *Reactions of two or more substrates*
465 *"Ping-pong" mechanisms*
466 *Isomechanisms*
466 *Dead-end complexes*
466 *Handling rate equations for complex mechanisms*
467 *The rapid equilibrium assumption*
467 *Isotope exchange at equilibrium*
468 5. Kinetics of Rapid Reactions
468 *Flowing substrates together*
468 *Observing relaxation*
468 *Rapid photometric methods*
468 *Some results*
469 6. Cryoenzymology
469 7. The Effect of pH on Enzymatic Action
471 **B. Inhibition and Activation of Enzymes**
471 1. Competitive Inhibitors
473 2. Noncompetitive Inhibition and Activation; Allosteric Sites
475 3. Inhibitors in the Study of Mechanisms
475 4. Allosteric Effectors in the Regulation of Enzyme Activity
477 5. Irreversible Inhibition of Enzymes
478 **C. The Specificity of Enzymatic Action**
478 1. Complementarity of Substrate and Enzyme Surfaces
478 2. Stereospecificity and Prochiral Centers
480 *Stereochemical numbering*
480 *Trigonal prochiral centers*
481 3. Induced Fit and Conformational Changes
481 4. Specificity and $k_{cat}$
482 5. Proofreading
482 **D. Mechanisms of Catalysis**
482 1. The "Transition State"
483 *Quantitative transition state theory*
484 *Transition state inhibitors*
484 *Describing the transition state*
486 *Getting to the transition state*
486 2. Microscopic Reversibility
486 3. Acid and Base Catalysis
486 *Acidic and basic groups in enzymes*
487 *Acid–base catalysis of mutarotation*
487 *General base and general acid catalysis*
490 *The Brönsted relationships*
490 *Concerted acid–base or tautomeric catalysis*
491 4. Hydrogen Bonding and the Transfer of Protons within Active Sites
491 *Ultra–fast proton transfer*
492 *Proton transfer rates*
493 *Marcus theory*
493 *Diffusion-controlled dissociation of protons*
493 *Coupled proton transfers*
493 *Unusually strong hydrogen bonds*
494 5. Covalent Catalysis
494 6. Proximity and Orientation
495 7. The Microenvironment
496 8. Strain and Distortion
497 9. Why Oligomeric Enzymes?
497 10. Summary
497 **E. Classification of Enzymes**

498 **References**
501 **Study Questions**

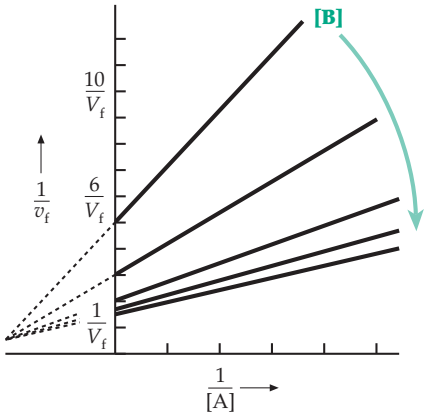
**Boxes**
456 Box 9-A A Historical Note on Enzymes
470 Box 9-B Growth Rates of Cells
473 Box 9-C The Sulfonamides as Antimetabolites
479 Box 9-D Receptors, Agonists, and Antagonists
485 Box 9-E Chorismate Mutase
488 Box 9-F Immunophilins as Rotamases

# Enzymes: The Catalysts of Cells

# 9




Random Bi Bi Mechanism




Most of the machinery of living cells is made of enzymes. Thousands of them have been extracted from cells and have been purified and crystallized. Many others are recognized only by their catalytic action and have not yet been isolated in pure form. Most enzymes are soluble globular proteins but an increasing number of RNA molecules are also being recognized as enzymes. Many structural proteins of the cell also act as catalysts. For example, the muscle proteins actin and myosin together catalyze the hydrolysis of ATP and link the hydrolysis to movement (Chapter 19). Catalysis is one of the most fundamental characteristics of life.

How do we recognize that a protein or RNA molecule *is* an enzyme? The answer is that *enzymes are recognized primarily by their ability to catalyze a chemical reaction.* For this reason, an everyday operation for many biochemists is the measurement of the catalytic activity of enzymes. Only by measuring the rates of the catalyzed reactions carefully and quantitatively has it been possible to isolate and purify these remarkable molecules.

Since the beginning of biochemical investigation enzymes have held a special fascination for chemists and biologists. How can these easily destroyed substances catalyze reactions with such speed and without formation of significant quantities of side products? Some enzymes increase the velocity of a single chemical reaction of a specific compound by a factor of as much as $10^{10}$. How can a protein do this? In this chapter we'll consider both ways of measuring enzymatic activity and basic mechanisms of catalysis.

## A. Information from Kinetics

The quantitative study of catalysis by enzymes, i.e., the study of enzyme kinetics, is a highly developed branch of biochemistry. It is one of our most important means of learning about the mechanisms of catalysis at the active sites of enzymes.[1–13a] By determining **rate constants *k*** under a variety of conditions we can learn just how fast an enzyme can act, how tightly it binds its substrates to form the enzyme-substrate (ES) complexes essential to catalysis, how specific it is with respect to substrate structure, and how it is affected by compounds that inhibit or activate the catalysis.

### 1. Measuring the Speed of an Enzymatic Reaction

A major goal in kinetic studies is to establish a **rate equation** which describes the velocity of a reaction in terms of **kinetic constants** and other experimentally measurable parameters. To measure the velocity of any chemical reaction precisely we must start the reaction at a definite time by rapidly mixing together the two or more reactants. Then, while keeping the mixture at an accurately constant temperature and pH, we must measure the concentration of a reactant or product after a fixed time interval, or at various times. No end of ingenuity has gone into devising ways of doing this for particular enzymes. Whatever the procedure, the information we must obtain is the **rate** at which some concentration changes with time. We can then construct a **progress curve** showing the decrease in concentration [S] of the reactant (**substrate**) or the increase in the concentration of **product** [P] with time (Fig. 9-1).

## BOX 9-A A HISTORICAL NOTE ON ENZYMES

While the earliest physiologists postulated a "vital force" to explain the chemical reactions of cells, the existence of biochemical substances promoting reactions outside of the body was recognized at least by the early 1600s. However, the role of yeast in fermentation was still unknown and it was thought that both alcoholic fermentation and animal digestion were caused by unknown substances called **ferments**. In 1752, Reamur demonstrated the solvent power of the gastric juice of birds and by 1783 Spallanzani had extended the studies to humans and other species.[a] In 1836 Schwann isolated the enzyme **pepsin**[b] from gastric juice.

In the same year Jacob Berzelius introduced the concept of catalysis, which he developed as a result of studies of the effects of acids and bases in promoting the hydrolysis of starch and of the effects of metals on the decomposition of hydrogen peroxide. Berzelius proposed the term **catalyst** from the Greek "katalysis," meaning "dissolution." Although he had been concerned primarily with inorganic catalysts, Berzelius recognized that a natural catalyst, an amylase that causes the hydrolysis of starch, had already been isolated from germinating barley in 1833 and that "in living plants and animals thousands of catalytic processes take place."[c] Some chemists and physiologists accepted Berzelius' concept of biochemical catalysis immediately, but many did not. The matter was complicated by Pasteur's discovery that yeast cells were the causative agent of alcoholic fermentation. Only after Edward Buchner's reports in 1897 that a juice formed by grinding yeast with sand and filtering could still ferment sugar (see Chapters 15 and 17) was the reality of enzymes in metabolism[d] generally accepted. The word **enzyme** (from the Greek "in yeast") was introduced earlier by F. W. Kuhne, professor of physiology in Heidelberg and a person who had accepted Berzelius' concept.[c]

[a] Richmond, C. (1986) *Trends Biochem. Sci.* **11**, 528–530
[b] Schwann, T. (1836) *Arch. Anat. Physiol.*, 90–138
[c] Hoffmann-Ostenhof, O. (1978) *Trends Biochem. Sci.* **3**, 186–187
[d] Buchner, E. (1897) *Ber. Deut. Chem. Ges.* **30**, 117–124

The velocity $v$ of an enzymatic reaction is defined as the rate at which a substrate disappears or at which a product is formed, the two being identical:

$$v = -d[\mathrm{S}]/dt = d[\mathrm{P}]/dt \qquad (9\text{-}1)$$

Under the steady-state conditions that usually apply (see p. 449), *the rate of increase of product will be the same as the rate of decrease of substrate.* The units of velocity are moles per liter per second ($\mathrm{M\ s^{-1}}$) or more traditionally in enzymology moles per liter per minute. We are interested in the **instantaneous velocity,** which at any time is given by the slope of the progress curve (Fig. 9-1). We usually want to measure the velocity immediately after the reaction is started to avoid the decrease in rate that comes from the depletion of substrate or accumulation of products. In some cases, as in the example given in Fig. 9-1, this is difficult to do with accuracy.

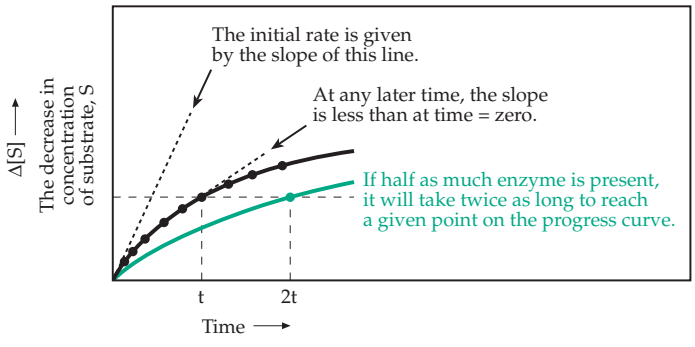



**Figure 9-1** The progress curve for an enzymatic reaction in which the substrate S is converted into products.

In some chemical reactions, which involve first-order processes, a logarithmic plot of the progress curve (log[S] vs $t$) gives a straight line so that the initial slope need not be determined. However, most **enzyme assays** give progress curves that remain nearly linear for only short periods of time. In these cases we need very sensitive methods for detecting products. These may involve colorimetric or fluorimetric measurements or the use of radioactive substrates. One of the most sensitive approaches is to arrange the assay so that a product of the reaction serves as a catalyst for another enzymatic process, thus amplifying the amount of final product to be measured.[14,15] Another approach is to measure velocities of reactions in a very small volume, e.g., 200 nanoliters, and to continuously separate and measure products using capillary electrophoresis and fluorescence detection.[16]

If the progress curve is not a straight line at the beginning of the reaction (0 time) and if the amount of compound that reacts in a fixed

interval of time is taken as the rate, an erroneous answer will be obtained. Sometimes an **integrated rate expression** that describes the time course of product formation can be used (see Eq. 9-22). Even when the progress curve is nonlinear as shown in Fig. 9-1, it is possible to estimate relative rates by noting that in most cases if one unit of enzyme yields a certain amount of product in time $t_1$, the same amount of product will be formed by $n$ units of enzyme in time $t_1/n$.

***First-order reactions.*** In many chemical reactions the rate of decrease of the concentration of a given reactant [A] is found experimentally to be directly proportional to the concentration of that reactant at any given time:

$$v = -d[\mathrm{A}]/dt = k[\mathrm{A}] \qquad (9\text{-}2)$$

Such a reaction is described as **first order** and the proportionality constant ***k*** is known as the **rate constant.** Such first-order kinetics is observed for **unimolecular processes** in which a molecule of A is converted into product P in a given time interval with a probability that does not depend on interaction with another molecule. An example is radioactive decay. Enzyme–substrate complexes often react by unimolecular processes. In other cases, a reaction is **pseudo-first order**; compound A actually reacts with a second molecule such as water, which is present in such excess that its concentration does not change during the experiment. Consequently, the velocity is apparently proportional only to [A].

A first-order rate constant $k$ has units of $s^{-1}$. When [A] = 1, $v = k$. Thus, $k$ is a measure of the speed in mol $l^{-1}$ $s^{-1}$ of the reaction of a substance at unit activity. As a first-order reaction proceeds, [A] decreases and at time $t$ is given by any one of the following three equivalent expressions (Eq. 9-3). These are obtained by integration of Eq. 9-2 in which $t_0$ is the time at which the reaction was started.

$$\begin{gathered}[\mathrm{A}] = A_0 e^{-kt} \\ \ln([\mathrm{A}_0]/[\mathrm{A}]) = kt \\ \log[\mathrm{A}_0] - \log[\mathrm{A}] = kt/2.303 \end{gathered} \qquad (9\text{-}3)$$

Equation 9-3 is the equation of exponential decay. A characteristic of exponential decay is that [A] is halved in a time that is independent of concentration. The **half-life** is $t_{1/2}$:

$$t_{1/2} = \ln 2/k = 0.693/k \qquad (9\text{-}4)$$

The **relaxation time** $\tau$ for A is defined by Eq. 9-5 and represents the time required for the concentration [A] to fall to $1/e$ (or ~0.37) of its initial value.

$$\tau = 1/\mathrm{k} = \mathrm{t}_{1/2}/\ln 2 \qquad (9\text{-}5)$$

***Turnover numbers and units of activity.*** When an enzyme is catalyzing product formation at the maximum possible rate, $V_{max}$, we can usually assume that the active site of every molecule of enzyme contains a substrate in the form of an intermediate enzyme–substrate complex ES and that this complex is being converted to products according to Eq. 9-6:

$$\begin{gathered} d[\mathrm{P}]/dt = V_{max} = k[\mathrm{ES}] = k[\mathrm{E}]_t \\ \text{or} \qquad V_{max} = k_{cat}[\mathrm{E}]_t \end{gathered} \qquad (9\text{-}6)$$

Here $E_t$ is the total enzyme, namely, the free enzyme E plus enzyme–substrate complex ES. The equation holds only at **substrate saturation,** that is, when the substrate concentration is high enough that essentially all of the enzyme has been converted into the intermediate ES. The process is first order in enzyme but is zero order in substrate. The rate constant $k$ is a measure of the speed at which the enzyme operates. When the concentration $[E]_t$ is given in moles per liter of *active sites* (actual molar concentration multiplied by the number of active sites per mole) the constant $k$ is known as the **turnover number**, the **molecular activity**, or $\boldsymbol{k_{cat}}$. The symbol $k_{cat}$ is also used in place of $k$ in Eq. 9-6 for complex rate expressions in which $k_{cat}$ cannot represent a single rate constant but is an algebraic expression that contains a number of different constants.

Turnover numbers can be measured only when the concentration of the enzyme is known. Partly for this reason the activity of an enzyme is usually given as **specific activity,** the *units of activity per milligram of protein*. One **international unit** is the amount of enzyme that produces 1 µmol of product per minute under standard (usually optimal) conditions. The International Union of Biochemistry[17] has recommended a larger unit, the **katal** (kat), the amount of enzyme that converts one mol $s^{-1}$ of substrate to product.

1 kat = $6 \times 10^7$ international units
1 international unit = 16.67 nkat (nanokatals)

If the enzyme is pure and is saturated with substrate under the standard assay conditions, the following relationships hold.

$$\begin{aligned} \text{Turnover No.} &= \text{katals/mol of active sites} \\ &= [\text{nkat/mg}] \times M_r \times 10^{-6}/n \\ &= [\text{international units/mg}] \times M_r \times 10^{-3}/60n \end{aligned}$$

Here $M_r$ is the relative molecular mass of the enzyme and $n$ is the number of active sites per molecule. Since the activity of an enzyme is dependent on both temperature and pH, these variables must be specified.

Turnover numbers of enzymes vary from <1 to ~ $10^6$ $s^{-1}$. Trypsin, chymotrypsin, and many intracellular

enzymes have turnover numbers of ~$10^2$ $s^{-1}$. The fastest enzymes, which include **catalase** (Chapter 16), **carbonic anhydrase** (Chapter 13), and **$\Delta^5$-3-oxosteroid isomerase** (Chapter 13), have maximum turnover numbers of 2 x $10^5$ $s^{-1}$ or more. Compare these reaction rates with those of a typical organic synthesis in the laboratory. A reaction mixture must often be heated for hours ($k < 10^{-3}$ $s^{-1}$). Many enzymes accelerate rates by factors of greater than $10^6$ over those observed in the absence of an enzyme at a comparable temperature and pH. Enzymes often bring two or more substrates together, binding them at a specific location in their active sites. Because of this, rapid reactions can be catalyzed even when the reactants are present in low concentrations.

***Second-order reactions.*** For a chemical reaction to occur between two molecules, A and B (Eq. 9-7), they must meet and collide.

$$A + B \xrightarrow{k_2} P \qquad (9\text{-}7)$$

The velocity of such a **second-order** process is characterized by a **bimolecular rate constant** $k_2$ and is proportional to the product of the concentrations of A *and* B:

$$v = k_2[A]\,[B] \qquad (9\text{-}8)$$

The units of $k_2$ are $M^{-1}$ $s^{-1}$. If [B] is present at unit activity, the rate is $k_2$[A], a quantity with units of $s^{-1}$. We can see that the bimolecular, or second-order, rate constant for reaction of A with B may be compared with first-order constants when the second reactant B is present at unit activity. In many real situations, reactant B is present in large excess and in a virtually constant concentration. The reaction is pseudo-first order and the experimentally observed rate constant $k_2$[B] is an *apparent first-order rate constant*. The bimolecular rate constant $k_2$ can be obtained by dividing the apparent constant by [B].

***Reversible chemical reactions.*** In any reversible process, we must consider rate constants for both the forward and the reverse reactions. At equilibrium a reaction proceeds in the forward direction at exactly the same velocity as in the reverse reaction so that no change occurs. For this reason there is always a relationship between the equilibrium constant and the rate constants. For Eq. 9-9, $k_1$ is the bimolecular rate constant

$$A + B \underset{k_2}{\overset{k_1}{\rightleftharpoons}} P$$

$$\text{Equilibrium constant} = K = \frac{[P]}{[A][B]} = \frac{k_1}{k_2} \qquad (9\text{-}9)$$

for the forward reaction and $k_2$ the unimolecular rate constant for the reverse reaction. The equilibrium constant $K$ can easily be shown, from Eq. 9-2 and 9-8, to equal $k_1/k_2$ for the reaction of Eq. 9-9.

The student should be aware that in kinetic equations rate constants are usually numbered consecutively via subscripts and that the subscripts do not imply anything about the molecularity. The system which is used here employs odd-numbered constants for steps in the forward direction and even-numbered constants for steps in the reverse direction. However, *many authors number the steps in the forward direction consecutively and those in the reverse direction with corresponding negative subscripts.*

What relationships exist between experimentally observable rates and $k_1$ and $k_2$ for a reversible reaction? Consider first the simplest case (Eq. 9-10):

$$A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} B \qquad (9\text{-}10)$$

If pure A is placed in a solution, its concentration will decrease until it reaches an equilibrium with the B which has been formed. It is easy to show that in this case [A] does not decay exponentially but [A] – $[A]_{equil}$ does. If log([A] – $[A]_{equil}$) is plotted against time a first-order rate constant $k$, characteristic of *the rate of approach to equilibrium*, will be obtained. Its relationship to $k_1$and $k_2$ is given by Eq. 9-11.

$$k = k_1 + k_2 \qquad (9\text{-}11)$$

The relaxation time $\tau$ for approach to equilibrium can be expressed as follows:

$$\tau^{-1} = k = \tau_1{}^{-1} + \tau_2{}^{-1} = k_1 + k_2 \qquad (9\text{-}12)$$

and for the more complex case of Eq. 9-9 as

$$k = \tau^{-1} = k_1([A]_e + [B]_e) + k_2 \qquad (9\text{-}13)$$

where $[A]_e$ and $[B]_e$ are the equilibrium concentrations of A and B.[18]

## 2. Formation and Reaction of Enzyme–Substrate Complexes

An abundance of evidence indicates that the first step in enzymatic catalysis is the combining of the enzyme and substrate reversibly to form a complex, ES (Eq. 9-14). Formation of the complex is normally reversible. ES can either break up to form enzyme and substrate again or it can undergo conversion to a product or products, often by a unimolecular process. Three rate constants are needed to describe this system for a reaction that is irreversible overall. A complete

description of the kinetic behavior is fairly involved.

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P \tag{9-14}$$

For example, the kinetics may be different within cells, where molar concentrations of enzymes often exceed those of substrate, than in the laboratory. In most laboratory experiments the enzyme is present at an extremely low concentration (e.g., $10^{-8}$ M) while the substrate is present in large excess. Under these circumstances the **steady-state approximation** can be used. For this approximation *the rate of formation of ES from free enzyme and substrate is assumed to be exactly balanced by the rate of conversion of ES on to P*. That is, for a relatively short time during the duration of the experimental measurement of velocity, the concentration of ES remains essentially constant. To be more precise, the steady-state criterion is met if the absolute rate of change of a concentration of a transient intermediate is very small compared to that of the reactants and products.[19]

The **Michaelis–Menten equation** (Eq. 9-15) describes the initial reaction rate of a single substrate with an enzyme under steady-state conditions.

$$v = \frac{V_{max}}{1 + K_m/[S]} = \frac{V_{max}[S]}{K_m + [S]}$$
$$\text{where } K_m = (k_2 + k_3)/k_1 \tag{9-15}$$

This can be rearranged as follows:

$$\frac{V_{max}}{v} = 1 + K_m/[S] \tag{9-16}$$

Equation 9-15 provides a relationship between the velocity observed at a particular substrate concentration and the maximum velocity that would be achieved at infinite substrate concentration. The quantities $V_{max}$ and $K_m$ are often referred to as the **kinetic parameters** of an enzyme and their determination is an important part of the characterization of an enzyme. Equation 9-15 can be derived by setting the rate of formation of the ES complex ($k1$[E][S]) in the steady state equal to its rate of breakdown, ([$k_2$ + $k_3$][ES]). Rearranging and substituting $K_m$, as defined in Eq. 9-15, we obtain Eq. 9-17.

$$[E][S] = \frac{(k_2 + k_3)}{k_1}[ES] = K_m[ES] \tag{9-17}$$

Using this equation, together with a mass balance relationship ([E] = $[E]_t$ – [ES]), we can solve for $[ES]/[E]_t$, the fraction of enzyme combined as enzyme–substrate complex (Eq. 9-18).

$$\frac{[ES]}{[E]_t} = \frac{[S]}{K_m + [S]} \tag{9-18}$$

The maximum velocity, $V_{max} = k_3[E]_t$, is attained only when all of the enzyme is converted into ES. Under other conditions $v = k_3$[ES] and Eq. 9-19 holds.

$$[ES] / [E]_t = v / V_{max} \tag{9-19}$$

Substituting from Eq. 9-19 into Eq. 9-18 gives the Michaelis–Menten equation (Eq. 9-15).

In many cases, the rate at which ES is converted back to free E and S is much greater than the rate of conversion of ES to products ($k_2 \gg k_3$). In such cases $K_m$ equals $k_2/k_1$, the dissociation constant for breakdown of ES to free enzyme and substrate (sometimes called $K_s$). Thus, *$K_m$ sometimes has a close inverse relationship to the strength of binding of substrate to enzyme.* In such cases, $1/K_m$ is a measure of the affinity of the substrate for its binding site on the enzyme and for a series of different substrates acted on by the same enzyme. The more tightly bound substrates have the lower values of $K_m$. But beware! The condition that $k_3$ is negligible compared to $k_2$ may be met with some (poorer) substrates but may not always be met by others. From Eq. 9-15 we see that $K_m$ is always greater than or equal to $K_s$, but it may be less than $K_s$ for more complex mechanisms.

Figure 9-2 shows a plot of velocity against substrate concentration as given by Eq. 9-15. The position of $V_{max}$ on the ordinate is marked, but it should be clear that the experimental velocity ($v$) can never attain $V_{max}$ unless [S] is very high relative to $K_m$. The value of $v$ approaches $V_{max}$ asymptotically. Since $K_m$ is defined as the value of [S] at which $v = V_{max}/2$, its value can be estimated from Fig. 9-2. However, $K_m$ cannot be determined reliably because of the difficulty of establishing the value of $V_{max}$ from a plot of this type. Notice that the curve of Figure 9-2 is identical in form to the saturation curve for reversible binding shown in Fig. 7-1.

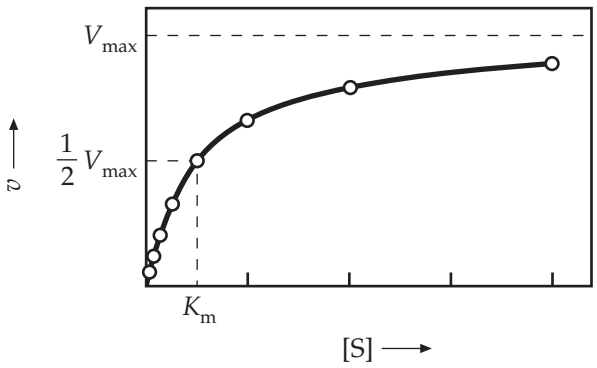



**Figure 9-2** Plot of observed velocity $v$ vs substrate concentration [S] for an enzyme-catalyzed reaction.

***Linear forms for rate equations.*** To obtain $K_m$ and $V_{max}$ from experimental rate data, Eq. 9-15 can be transformed by algebraic rearrangement into one of several linear forms. The popular **double-reciprocal** or **Lineweaver–Burk** plot of $1/v$ against $1/[S]$ (Fig. 9-3) is described by Eq. 9-20. The values of $K_m/V_{max}$ and $1/V_{max}$ can be evaluated from the slope and intercept, respectively, of this straight line plot.

$$\frac{1}{v} = \frac{1}{V_{max}} + \frac{K_m}{V_{max}}\frac{1}{[S]} \qquad (9\text{-}20)$$

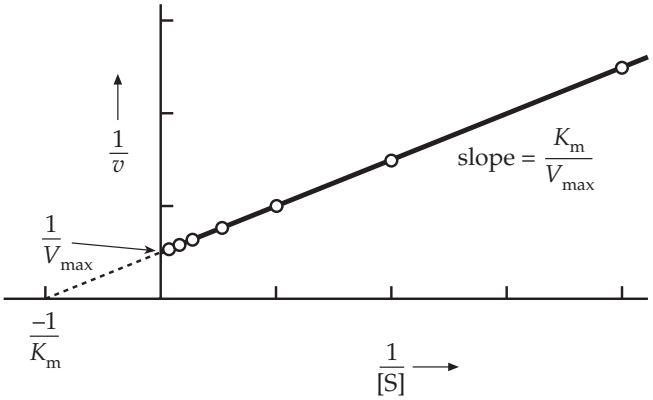



**Figure 9-3** Double-reciprocal or Lineweaver–Burk plot of $1/v$ vs $1/[S]$. The intercept on the vertical axis gives $1/V_{max}$ and the slope gives $K_m/V_{max}$. The intercept on the horizontal axis equals $-1/K_m$.

Another linear plot, the **Eadie–Hofstee plot**, is that of $v/[S]$ vs $v$ (Fig. 9-4). It is related to the Scatchard plot (Fig. 7-3) and is fitted by Eq. 9-21.

$$\frac{v}{[S]} = \frac{V_{max}}{K_m} - v\frac{1}{K_m} \qquad (9\text{-}21)$$

While the point for $[S] = 0$ and $v = 0$ cannot be plotted, the ratio $v/[S]$ approaches $V_{max}/K_m$ as $v$ approaches zero. Notice the distribution of the points in Fig. 9-4. Substrate concentrations were chosen such that the increase in velocity from point to point is more or less constant, a desirable experimental situation. The points on the Eadie–Hofstee plot are also nearly evenly distributed, but those of the Lineweaver–Burk plot are compressed at one end. (However, if the substrate concentrations for successive points are selected in the ratios 1, 1/3, 1/5, 1/7, and 1/9, the spacing will be uniform on the Lineweaver–Burk plot.) A second advantage of the Eadie–Hofstee plot is that the entire range of possible substrate concentration from near zero to infinity can be fitted onto a single plot.

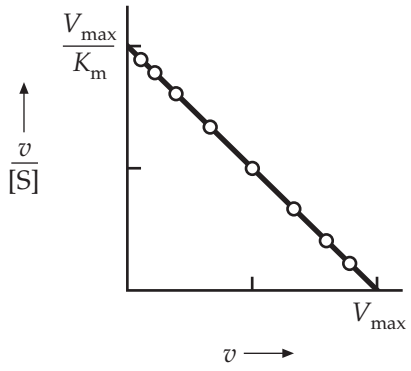



**Figure 9-4** The Eadie–Hofstee plot of $v/[S]$ against $v$. The slope is $-1/K_m$; the intercept on the vertical axis is $V_{max}/K_m$ and that on the horizontal axis is $V_{max}$.

As pointed out at the beginning of Section A, depletion of substrate with time lowers [S] from its initial value to some extent. This can have an especially adverse effect on points obtained at low [S], e.g., points on the right side of Fig. 9-3 or on the left side of Fig. 9-4. Equations 9-20 and 9-21 will be more precise if the average substrate concentration over the time period of the assay, rather than that at zero time,[20] is used.

***Nonlinear equations and integrated rate equations.*** In discussions of the *control* of metabolism through regulation of enzymatic activity, it is often better to plot $v$ against log [S] as in Fig. 9-5. This plot also has the virture that the entire range of attainable substrate concentrations can be plotted on one piece of paper if the point for [S] = 0 (at minus infinity) is omitted. The same scale can be used for all enzymes. The plot is S-shaped, both for simple cases that are represented by Eq. 9-15 and for enzymes that bind substrate cooperatively (Section B,5). Thus, the classification or "hyperbolic" vs "sigmoidal" is lost. However, the degree of cooperativity can be directly measured from the midpoint slope of the curve. A disadvantage is that it is awkward to measure $V_{max}$ from a plot of this type, and it may be preferable to obtain it from a linear plot (Figs. 9-3 or 9-4). Alternatively, computer-assisted methods can be used to obtain both $K_m$ and $V_{max}$ and to fit a curve to the experimental points as in Fig. 9-5.

An attractive alternative to the use of initial velocities and linear plots is to measure the kinetic parameters $V_{max}$ and $K_m$ using points all along the progress curve (Fig. 9-1). Various procedures for doing this have been devised.[21,22] For example, the integrated form of Eq. 9-15 can be given as Eq. 9-22 or 9-23.

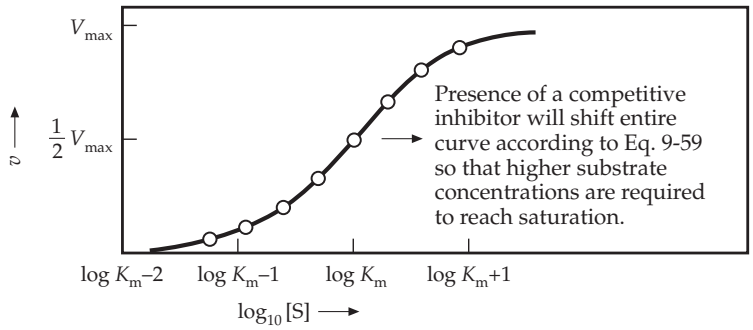



**Figure 9-5** Plot of $v$ against log [S] for an enzyme-catalyzed reaction.

$$V_{max}t = [P] + K_m \ln \frac{[S]_0}{[S]_0 - [P]} \qquad (9\text{-}22)$$

$$V_{max}\, t = [S]_0 - [S] + K_m \ln ([S]_0 / [S]) \qquad (9\text{-}23)$$

Here $[S]_0$ is the initial substrate concentration at $t = 0$ and [S] and [P] are the substrate and product concentrations at any later time. In one type of linear plot,[21] based on Eq. 9-23, $([S]_0 - [S]) / t$ is plotted against $1/t$ ln $([S]_0 / [S])$, the slope of the plot being $-K_m$ and the intercepts being $V_{max}$ and $V_{max} / K_m$. Computer programs are available for analysis of enzymatic progress curves, even for complex mechanisms.[23,24]

***Kinetics with high enzyme concentrations.*** Laboratory studies of the kinetics of purified enzymes are usually conducted with enzyme concentrations of $10^{-7}$ to $10^{-10}$ M, but within cells enzyme concentrations are probably often in the range $10^{-6}$ to $10^{-5}$ M,[25] which may be higher than the concentrations of the substrates upon which the enzymes act. Be cautious in drawing conclusions about kinetics under such circumstances! Methods have been devised for handling kinetic data when the concentration of enzyme is greater than $K_m$, a condition that can lead to intolerably high errors if the usual equations are applied.[26,27]

## 3. Diffusion and the Rate of Encounter of an Enzyme with Substrate

What determines the value of $k_1$ in Eq. 9-14? This rate constant represents the process by which the substrate and enzyme find each other, become mutually oriented, and bind to form ES. If orientation and binding are rapid enough, the rate will be determined by the speed with which the molecules can come together by diffusion. Large molecules in solution are free to travel for only a tiny fraction of their diameter as a result of their frequent collision with solvent molecules. The result is visible in the **Brownian movement** of microscopic particles suspended in a fluid. If an individual particle is followed, it is seen to undergo a "random walk," moving in first one direction then another. Albert Einstein showed that if the distances transversed by such particles in a given time $\Delta t$ are measured, the mean square of these $\Delta x$ values $\Delta^2$ can be related by Eq. 9-24 to the diffusion constant $D$ (which is usually given in units of $cm^2 s^{-1}$).

$$\Delta^2 = 2D\,\Delta t \qquad (9\text{-}24)$$

For molecules, the Brownian movement cannot be observed directly but the diffusion constant can be measured, for example, by observing the rate of spreading of a boundary between two different concentrations of the substance.[28] The diffusion constant for $^1H^2HO$ (HDO) in $H_2O$ at 25°C is $2.27 \times 10^{-5}$ cm$^2$ s$^{-1}$, and the values for the ions $K^+$ and $Cl^-$ are about the same.[29] For many small molecules $D$ is approximately $10^{-5}$ cm$^2$ s$^{-1}$, and the value decreases as the size of the molecule increases. Thus, for the 13.7-kDa ribonuclease, $D = 1.1 \times 10^{-6}$ cm$^2$ s$^{-1}$, and for the 500-kDa myosin, $D$ is ~ $1 \times 10^{-7}$ cm$^2$ s$^{-1}$. For the spherical particle the **Stokes–Einstein equation** (Eq. 9-25) can be used to relate the diffusion coefficient to the radius, the coefficient of viscosity η, the Boltzmann constant $k_B$, and the temperature (Kelvin) $T$.

$$D = \frac{k_B T}{6\pi\eta r} \qquad (9\text{-}25)$$

To estimate the rate constant for a reaction that is controlled strictly by the frequency of collisions of particles, we must ask how many times per second one of a number $n$ of particles will be hit by another of the particles as a result of Brownian movement. The problem was analyzed in 1917 by Smoluchowski,[30,31] who considered the rate at which a particle B diffuses toward a second particle A and disappears when the two collide. Using Fick's law of diffusion, he concluded that the number of **encounters** per milliliter per second was given by Eq. 9-26.

$$\text{Number of encounters/ml s}^{-1} = 4\,\pi\,(D_A + D_B)\,(r_A + r_B)\,n_A\,n_B \qquad (9\text{-}26)$$

Here $D_A$ and $D_B$ are the two diffusion coefficients, $r_A$ and $r_B$ are the radii (in Å) of the two particles, and $n_A$ and $n_B$ are the numbers of particles per milliliter. The number of encounters per liter per second is $(4\pi / 1000)\,(D_A + D_B)\,(r_A + r_B)\,N^2\,[A][B]$, where $N$ is Avogadro's number. Dividing this frequency by $N$ gives the rate of collision $v$ in M s$^{-1}$, a velocity that can be equated (Eq. 9-27) with $k_D\,[A]\,[B]$, where $k_D$ is a second-order rate constant

which determines the **diffusion controlled limit**, that is, the maximum possible rate of a reaction.

$$v = \frac{4\pi}{1000}(D_A + D_B)(r_A + r_B)[A][B]N = k_D[A][B] \quad (9\text{-}27)$$

This equation can be rearranged to Eq. 9-28.

$$k_D = \frac{4\pi N}{1000}(D_A + D_B)(r_A + r_B)\mathrm{M}^{-1}\,\mathrm{s}^{-1} \quad (9\text{-}28)$$

While this equation is thought to overestimate the diffusion-limited rate constant slightly, it is a good approximation. If the diffusing particles are approximately spherical, diffusion constants $D_A$ and $D_B$ can be calculated from Eq. 9-25, and Eq. 9-28 becomes Eq. 9-29.

$$k_D = \frac{2RT}{3000\eta}\left(2 + \frac{r_A}{r_B} + \frac{r_B}{r_A}\right) \quad (9\text{-}29)$$

The value of $k_D$ does not vary greatly as the ratio of radii $r_A / r_B$ is changed, and in most cases it may be assumed that $r_A / r_B \approx 1$, in which case the equation simplifies (Eq. 9-30).

$$k_D \approx \frac{8RT}{3000\eta}\mathrm{M}^{-1}\,\mathrm{s}^{-1} \quad (9\text{-}30)$$

For water at 25°C, the coefficient of viscosity η is ~0.01 poise (1 poise = $10^{-5}$ newton $cm^{-2}$) which leads, according to Eq. 9-30, to $k_D \sim 0.7 \times 10^{10}$ $M^{-1}$ $s^{-1}$. It was shown by Debye[32] that this rate constant must be multiplied by a correction factor when charged particles rather than uncharged spheres diffuse together. This factor may be of the order of 5–10 for a substrate and enzyme carrying two or three charges and may act to either increase or decrease reaction rates. Since the viscosity of cytoplasm is 0.02–0.03 poise,[33] $k_D$ will be reduced correspondingly.

A simple alternative derivation of Eq. 9-30 was developed by Dexter French,[34] who also provided the author with most of this discussion of encounter theory. Consider a small element of volume $\Delta V$ swept out by a particle as it moves through the solution for a distance equal to its own radius. This element of volume will equal $\pi r^3$:

$$\Delta V = \pi r^2 \times r = \pi r^3 \quad (9\text{-}31)$$

It will be swept out in a time $\Delta t$ which can be calculated from Eq. 9-24 as $r^2 / 2D$. Substituting the value of $D$ given by Eq. 9-25, we obtain for $\Delta t$:

$$\Delta t = \frac{3\pi\eta r^3}{k_B T} \quad (9\text{-}32)$$

Division of Eq. 9-31 by Eq. 9-32 gives the volume swept out per second by one particle:

$$\Delta V / \Delta t = k_B T / 3\eta \ \mathrm{cm}^3\,\mathrm{s}^{-1} \quad (9\text{-}33)$$

Since the collision radius for two particles of equal size is two times the particle radius, the effective volume swept out will be four times that given by Eq. 9-33. Since both particles are diffusing, the effective diffusion constant will be twice that used in obtaining Eq. 9-28. Thus, the effective volume swept out by the particle in a second will be eight times that given by Eq. 9-33. The volume swept out by one mole of particles is equal to $k_D$ (recall that the second-order rate constant has dimensions of liter $mol^{-1}$ $s^{-1}$). Thus, when converted to a moles per liter basis and multiplied by 8, Eq. 9-33 should (and does) become identical with the Smoluchowski equation (Eq. 9-30).

The volume given by Eq. 9-33 is about $1.4 \times 10^{-11}$ $cm^3$, which could be represented approximately by a cube 2.4 μm on a side. If we compare this volume with that of a cell (Table 1-2) or of an organelle, we see that in one second an enzyme molecule will sweep out a large fraction of the volume of a small cell, mitochondrion, chloroplast, etc.

***The "cage effect" and rotation of molecules.*** It is of interest to compare the bimolecular rate constant for encounters calculated by the Smoluchowski theory (Eq. 9-29) with the corresponding bimolecular rate constant for molecular collisions given by the kinetic theory of gases (Eq. 9-34).

$$k(\mathrm{collision}) = \frac{N(r_A + r_B)^2}{1000}\left[8\pi k_B T\left(\frac{1}{m_A} + \frac{1}{m_B}\right)\right]^{1/2} \quad (9\text{-}34)$$

Here $m_A$ and $m_B$ are the masses of the two particles. This rate constant is also relatively independent of molecular size and for spheres varies from (4 to 11) x $10^{11}$ $M^{-1}$ $s^{-1}$, over an order of magnitude greater than the rate constant for encounters. In a solution, molecules still collide at about the same rate as in a gas so that *100 to 200 collisions occur between two particles for each encounter*. However, during the time of the single encounter, the particles are together in a solvent "cage." While in this solvent cage, both substrate and enzyme molecules undergo random rotational motions. Successive collisions bring them together in different orientations, one of which is likely to lead to a sufficiently close match of complementary surfaces (of substrate and binding site) so that formation of a "productive" ES complex takes place.

Molecular rotation in a solution is described quantitatively by diffusion laws (analogous to Fick's laws)

for which a **rotary diffusion constant** θ is defined.[35,36] Consider a group of molecules all oriented the same way initially, then undergoing rotary diffusion until their orientations become random. If we measure the orientation of each molecule by an angle we see that initially the value of cos α is 1 but that when the angles become random the mean value of cos α averaged over all molecules is zero. The rotary relaxation time τ is the time required for the mean value of cos α to fall to


$1/e$ (which occurs when $\alpha = 68.5°$). For a sphere θ is given by Eq. 9-35.

$$\theta = \tfrac{1}{2}\tau = \frac{k_B T}{8\pi\eta r^3} \qquad (9\text{-}35)$$

Ellipsoidal or rod-shaped molecules have two different rotary diffusion constants while, if the dimensions of the molecules are different along all three axes, three constants must be specified.[36]

From Eq. 9-35 we can calculate (if $\eta = 0.01$ poise) the following values for a small spherical molecule (substrate) of ~1 nm length and for a spherical enzyme of 5 nm diameter:

$r = 0.5$ nm $\quad \theta \approx 1.3 \times 10^9\ s^{-1}$
$r = 2.5$ nm $\quad \theta \approx 1.0 \times 10^7\ s^{-1}$

We see that smaller molecules rotate much faster than large ones and that rotational relaxation times for small proteins are of the same order of magnitude as $k_D$ for diffusion-limited encounter. However, for very large molecules, especially long rods, the rotary relaxation time about a short axis may be a large fraction of a second.

***The rate of substrate binding.*** At very low substrate concentrations the Michaelis–Menten equation (Eq. 9-15) simplifies as follows:

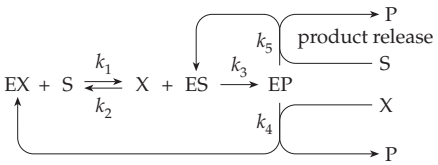
$$v = (V_{max}/K_m)[S] \qquad (9\text{-}36)$$

Since $V_{max} = k_{cat}[E]_t$ (Eq. 9-6) and at low [S] most of $[E]_t$ is free E, we obtain the following equation:[2]

$$v = (k_{cat}/K_m)[E][S] \qquad (9\text{-}37)$$

From this we see that $k_{cat}/K_m$ is *the apparent second-order rate constant for the reaction of free enzyme with substrate*. As such it cannot exceed the diffusion controlled limit $k_D$ of Eqs. 9-28 to 9-30 which falls in the range of $10^9 - 10^{11}\ M^{-1}\ s^{-1}$. Experimentally observed values of $k_{cat}/K_m$ are always less than this limit, indicating that a certain time is required for a substrate molecule to become oriented and seated in the active site.[37] However, for several real enzymes values of $k_{cat}/K_m$ of $10^7$ to $3 \times 10^8\ M^{-1}\ s^{-1}$ have been observed. For triose phosphate isomerase Albery and Knowles obtained a value of $4 \times 10^8\ M^{-1}\ s^{-1}$, so close to the diffusion controlled limit that these authors regard this enzyme as a *nearly perfect catalyst*, one that could not have evolved further because it is already catalyzing the reaction with substrate at almost the maximum velocity that is possible.[37–39]

***The displacement of bound ligands by substrate.*** Jenkins pointed out that in many instances a substrate must displace another ligand from the active site to form the ES complex.[40] For example, a binding site for an ionic substrate often already contains an ion, either of the product P or some other ion X which Jenkins calls a **substrate surrogate**. For this situation Eq. 9-14 must be replaced by the following set of equations:




(9-38)

Here $k_2$ and also $k_4$ and $k_5$, are second-order rate constants. The release of product, as determined by $k_4$ and $k_5$, may be rate-limiting. At zero time the reverse reactions may be ignored, and steady-state analysis shows that the Michaelis–Menten equation (Eq. 9-16b) will be replaced by Eq. 9-39. Here, D is a constant and A is also constant if X is present at a fixed concentration.

$$\frac{V_{max}}{v} = 1 + \frac{k_2[X] + k_3}{k_1[S]} + \frac{k_3(1/k_5 - 1/k_1)}{k_4[X]/k_5 + [S]}$$
$$= 1 + K_m/[S] + D/(A + [S]) \qquad (9\text{-}39)$$

When $k_5$, the rate constant for displacement of product by substrate, is very small, this equation simplifies to Eq. 9-40.

$$\frac{V_{max}}{v} = 1 + k_3/k_4[X] + (k_2[X] + k_3)/k_1[S]$$
$$= 1 + k_3/k_4[X] + K_m'/[S] \qquad (9\text{-}40)$$

At high concentrations of X the second term becomes negligible and the equation becomes identical to the Michaelis–Menten equation (Eq. 9-16) except that $K_m'$, the apparent Michaelis constant, now increases as [X]

increases (Eq. 9-41). This is exactly what is observed for competitive inhibitors where

$$K_m' = \frac{k_2[\mathrm{X}] + k_3}{k_1} \qquad (9\text{-}41)$$

Thus substrate surrogates at high concentrations are often competitive inhibitors. However, at low [X] the second term in Eq. 9-41 may be important and *X then serves as an activator*. Under many circumstances, Eq. 9-39 does not simplify further and a nonlinear relationship between $1/v$ and $1/[\mathrm{S}]$ is observed, with the shape of the curves being influenced by the values of constants A and D (see also Section B,4).[40]

## 4. Reversible Enzymatic Reactions

For some enzyme-catalyzed reactions the equilibrium lies far to one side. However, many other reactions are freely reversible. Since a catalyst promotes reactions in both directions, we must consider the action of an enzyme on the reverse reaction. Let us designate the maximum velocity in the forward direction as $V_f$ and that in the reverse direction as $V_r$. There will be a Michaelis constant for reaction of enzyme with product $K_{m\mathrm{P}}$, while $K_{m\mathrm{S}}$ will refer to the reaction with substrate.

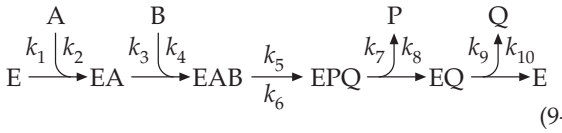
As in any other chemical reaction, there is a relationship between the rate constants for forward and reverse enzyme-catalyzed reactions and the equilibrium constant. This relationship, first derived by the British kineticist J. B. S. Haldane and proposed in his book *Enzymes*[41] in 1930, is known as the **Haldane relationship**. It is obtained by setting $v_f = v_r$ for the condition that product and substrate concentrations are those at equilibrium. For a single substrate–single product system it is given by Eq. 9-42.

$$K_{\mathrm{eq}} = V_f K_{m\mathrm{P}} / V_r K_{m\mathrm{S}} \qquad (9\text{-}42)$$

Because the Haldane relationship imposes constraints on the values of the velocity constants and Michaelis constants, it is of some value in understanding regulation of metabolism. Consider the case that the maximum forward velocity $V_f$ is high and that $K_{m\mathrm{S}}$ has a moderately low value (fairly strong binding of substrate). If the reaction is freely reversible ($K_{\mathrm{eq}} \sim 1$) and the velocity of the reverse reaction $V_r$ is about the same as that of the forward reaction, it will necessarily be true (see Eq. 9-42) that the product P will also be fairly tightly bound. If $V_r << V_f$, the value of $K_{m\mathrm{P}}$ will have to be much *lower* than that of $K_{m\mathrm{A}}$. In such a situation P will remain tightly bound to the enzyme and since $V_r$ is low, it will tend to clog the enzyme. Such **product inhibition** may sometimes slow down a whole pathway of metabolism. In such a case, the only way that an enzymatic sequence can keep going in the forward direction is for product P to be removed rapidly by a subsequent reaction with a second enzyme. The first enzyme may be thought of as possessing a kind of "one-way valve" that turns off the flow in a metabolic pathway when the concentration of its product rises.

***Reactions of two or more substrates.*** Enzymes frequently catalyze the reaction of two, three, or even more different molecules to give one, two, three, or more products. Sometimes all of the substrate molecules must be bound to an active site at the same time and are presumably lined up on the enzyme molecule in such a way that they can react in proper sequence. In other cases, the enzyme may transform molecule A to a product, and then cause the product to react with molecule B. The number of variations is enormous.[1,10,12]

The order in which two molecules A and B bind to an enzyme to form a complex EAB may be completely *random* or it may be *obligatorily ordered*. Both situations occur with real enzymes. Cleland introduced a widely used method of depicting the possibilities.[42] For example, Eq. 9-43 shows the reaction of A and B in an *ordered sequence* to form the complex EAB which is then isomerized to EPQ, the complex formed by binding the two products P and Q to the enzyme. The rate constant to the left of each vertical arrow or above each horizontal arrow refers to the reaction in the




(9-43)

forward direction as indicated by the arrow while the other constants (to the right or below the arrows) refer to the reverse reactions. The velocity in the forward direction for an enzyme with ordered binding is given by Eq. 9-44a,

$$v_f = \frac{V_f[\mathrm{A}][\mathrm{B}]}{K_{\mathrm{eqA}}K_{m\mathrm{B}} + K_{m\mathrm{B}}[\mathrm{A}] + K_{m\mathrm{A}}[\mathrm{B}] + [\mathrm{A}][\mathrm{B}]} \qquad (9\text{-}44a)$$

which may also be written in the reciprocal form (Eq. 9-44b):

$$\frac{1}{v_f} = \frac{1}{V_f}\left(1 + \frac{K_{m\mathrm{A}}}{[\mathrm{A}]} + \frac{K_{m\mathrm{B}}}{[\mathrm{B}]} + \frac{K_{\mathrm{eqA}}K_{m\mathrm{B}}}{[\mathrm{A}][\mathrm{B}]}\right) \qquad (9\text{-}44b)$$

An alternative form of this equation (Eq. 9-45), proposed by Dalziel,[43] is sometimes used.

$$\frac{[E]_t}{v_f} = \phi_0 + \frac{\phi_1}{[A]} + \frac{\phi_2}{[B]} + \frac{\phi_{12}}{[A][B]} \qquad (9\text{-}45)$$

The kinetic parameters of Eq. 9-44 are $V_f$, the maximum velocity in the forward direction, the two Michaelis constants, $K_{mB}$ and $K_{mA}$, and the equilibrium constant $K_{eqA}$, for reversible dissociation of the complex EA and which is equal to $k_2/k_1$. The relationship between the parameters of Eq. 9-44 ($K_m$'s, $V$'s, and $K_{eqA}$'s) and the rate constants $k_1 - k_{10}$ is not obvious. However, remember that the parameters are experimental quantities determined by measurements on the enzyme. Sometimes, but not always, it is possible to deduce some of the values of individual rate constants from the experimental parameters.

An equation similar to 9-44a can be written for the velocity $v_r$ of the reaction of P and Q. Also an equation can be written in similar form for $v_f - v_r$, i.e., the instantaneous velocity of the reaction in any mixture of all four components A, B, P, and Q.

The kinetic parameters of Eq. 9-44b are often obtained from experimental data through the use of reciprocal plots (Fig. 9-6). However, Eq. 9-44b is linear only if the concentration of one or the other of the substrates A and B is kept constant. For this reason a series of experiments is usually performed in which [A] is varied while [B] is held constant. Then [A] is held constant and [B] is varied. Each of these experiments leads to a family of lines (Fig. 9-6A) whose slopes and intercepts are measured. The slopes and intercepts of this family of curves are then plotted against the reciprocal of the second concentration, i.e., the one that was held fixed.

From a set of these **secondary plots**, $V_f$ and one of the Michaelis constants can be determined (Fig. 9-6B and C). Using two sets of secondary plots, all of the constants of Eq. 9-44 may be established. Alternatively, a computer can be used to examine all of the data at once and to obtain the best values of the parameters. The latter approach is desirable because estimates of the standard deviations of the parameters can be obtained. However, the user must take care to ensure that the experimental errors are correctly estimated and are not simply estimates of how well the computer has fitted the points on the assumption that they contain no error.[23]

The meaning of the kinetic parameters may be slightly difficult to grasp. $V_f$ is the velocity that would be obtained if both [A] and [B] were high enough to saturate the enzyme. Each $K_m$ corresponds to that for a simple system in which the concentration of the second substrate is at a high, saturating value.

For the bimolecular reaction that we have considered, there are two Haldane relationships:

$$K_{eq} = \frac{V_f K_{mP} K_{dQ}}{V_r K_{dA} K_{mB}} = \left(\frac{V_f}{V_r}\right)^2 \frac{K_{dP} K_{mQ}}{K_{mA} K_{dB}} \qquad (9\text{-}46)$$

Of these, only the first is ordinarily used.

***"Ping-pong" mechanisms.*** A common type of mechanism that is especially prevalent for enzymes with tightly bound cofactors has been dubbed **ping-pong** because the enzyme alternates between

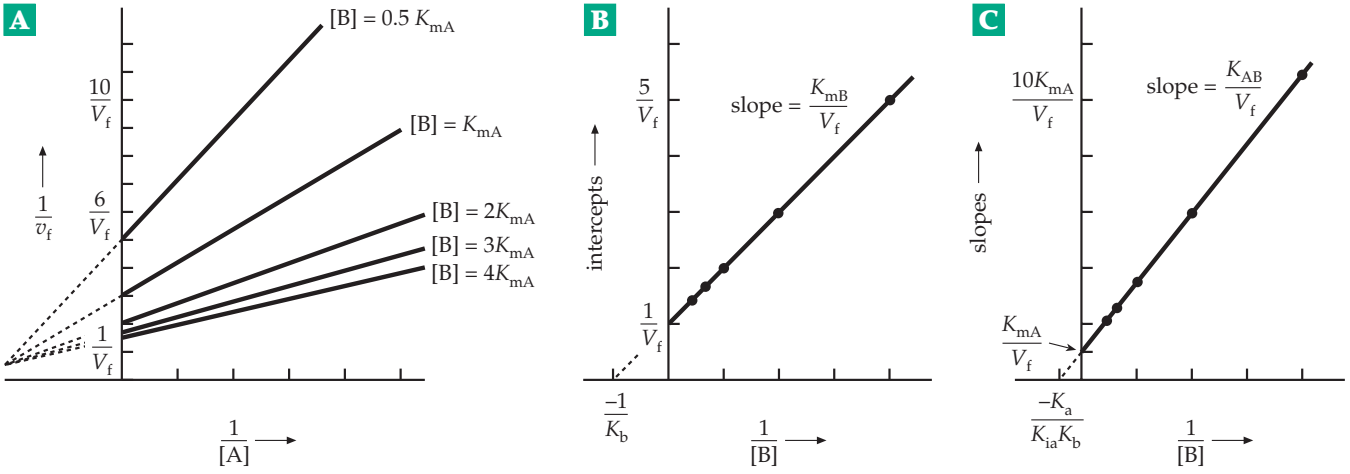
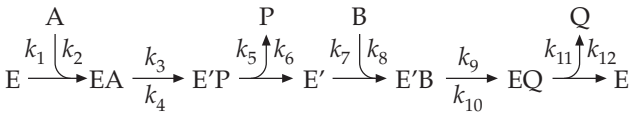



**Figure 9-6** Reciprocal plots used to analyze kinetics of two-substrate enzymes. (A) Plot of $1/v_f$ against $1/[A]$ for a series of different concentrations of the second substrate B. (B) A secondary plot in which the intercepts from graph A are plotted against $1/[B]$. (C) Secondary plot in which the slopes from graph A have been plotted against $1/[B]$. The figures have been drawn for the case that $K_{mA} = 10^{-3}$ M, $K_{mB} = 2\,K_{mA}$, and $K_{AB} = K_{eqA}K_{mB}$ (Eq. 9-46) $= K_{mA}/200$ and [A] and [B] are in units of moles per liter. Eadie–Hofstee plots of $v_f/[A]$ vs $v_f$ at constant [B] can also be used as the primary plots. The student can easily convert Eq. 9-44 to the proper form analogous to Eq. 9-21.

two forms E and E′ (Eq. 9-47). Substrate A reacts via complex EA to form E′, a modified enzyme that often




Ping-pong mechanism for reaction of A and B to form P and Q (9-47)

contains an altered coenzyme. At the same time A is changed to product P still bound to the enzyme. P dissociates leaving E′ which is then able to react with the second substrate B and to go through the second half of the cycle during which E′ is converted back to E. An example is provided by the aminotransferases (Eq. 14-25) in which the coenzyme pyridoxal phosphate is interconverted with pyridoxamine phosphate.

The rate equations of the ping-pong mechanism resemble that for the ordered bimolecular reaction (Eq. 9-44), but each has one less term (Eq. 9-48):

$$v_f = \frac{V_f[A][B]}{K_{mB}[A] + K_{mA}[B] + [A][B]} \qquad (9\text{-}48a)$$

or

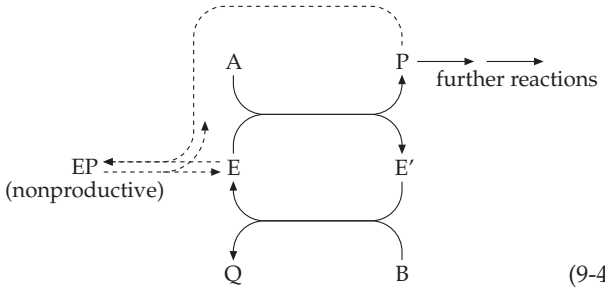
$$\frac{1}{v_f} = \frac{1}{V_f}\left(1 + \frac{K_{mA}}{[A]} + \frac{K_{mB}}{[B]}\right) \qquad (9\text{-}48b)$$

A diagnostic feature of the ping-pong mechanism is that the families of lines (in double reciprocal plots) obtained when one substrate is held constant while the other is varied no longer intersect as they do in Fig. 9-6A but are parallel. They must be truly parallel and the experimentalist must be aware that nearly parallel lines may sometimes be observed for sequential reactions. Thus, if $K_{eqA}$ of Eq. 9-44b is small enough the last term of that equation will be small compared to the other terms and the equation will be approximated by Eq. 9-48b. The reaction will appear to be ping-pong even though it is sequential and the reaction proceeds through the ternary complex EAB.

One less kinetic parameter can be obtained from an analysis of the data for a ping-pong mechanism than can be obtained for ordered reactions. Nevertheless, in Eq. 9-47, twelve rate constants are indicated. At least this many steps must be considered to describe the behavior of the enzyme. Not all of these constants can be determined from a study of steady-state kinetics, but they may be obtained in other ways.

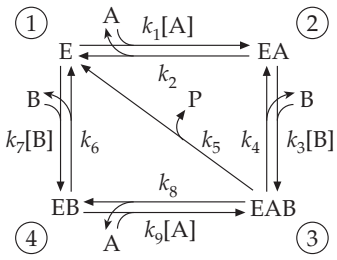
***Isomechanisms.*** A catalyst functions over and over again without being altered. However, during a single turnover it is likely to undergo a temporary change. For example, if an enzyme assists in removing a proton from a substrate, some functional group of the enzyme will be protonated during part of the catalytic cycle. This proton must dissociate before another catalytic cycle begins, and in some instances this dissociation may be a rate-limiting step. Northrup and coworkers developed the use of product inhibition to identify such isomechanisms.[44]

***Dead-end complexes.*** In the steady state of action of an enzyme with ping-pong kinetics, part of the enzyme is in form E and part in form E′. Ideally, E would have affinity only for A and Q, while E′ would have affinity only for B and P. However, in many real situations E also binds B and P weakly; similarly, E′ binds A and Q. This is because the products and reactants usually have structural features in common. The propensity of enzymes with ping-pong kinetics to form **dead-end complexes** (also called abortive complexes) may sometimes have a regulatory function. In Eq. 9-49 the reactions of Eq. 9-47 have been rearranged to depict a situation in which product P normally undergoes a sequence of further reactions. However, if P accumulates to a high enough concentration it can react reversibly to form a dead-end complex EP. This is an effective form of product inhibition which can be relieved only by a lowering of the concentration of P through its further metabolism.




(9-49)

***Handling rate equations for complex mechanisms.*** While steady-state rate equations can be derived easily for the simple cases discussed in the preceding sections, enzymes are often considerably more complex and the derivation of the correct rate equations can be extremely tedious. The **topological theory of graphs**, widely used in analysis of electrical networks, has been applied to both steady-state and nonsteady-state enzyme kinetics.[45–50] The method employs diagrams of the type shown in Eq. 9-50. Here




(9-50)

the reaction of an enzyme with two substrates A and B with a **random** order of binding is depicted. (In contrast, Eq. 9-43 shows the case of ordered binding of two substrates.) When complex EAB is formed, it can decompose to free enzyme and to the single product P. Each one of the **nodes**, which are numbered 1–4 in the diagram, corresponds to a single form of the enzyme. The appropriate first-order rate constant or apparent first-order constant is placed by each arrow. The methods provide easy rules for deriving from such a scheme the steady-state rate equation.

The importance of the simplified schematic methods is apparent when one considers that the steady-state rate equation for Eq. 9-50 would have 6 terms in the numerator and 12 terms in the denominator.[51] In the more complex case in which EAB breaks down to two products P and Q with a random order of release, the rate equation contains 672 terms in the denominator. In such cases it is worthwhile to enlist the help of a computer in deriving the equation.[24,52–54]

***The rapid equilibrium assumption.*** Rate equations for enzymes are often simplified if a single step, e.g., that of reaction of complex EAB to product in Eq. 9-50, is **rate limiting**.[54a] If it is assumed that all reaction steps preceding or following the rate-limiting step are at equilibrium, the equation for random binding with a two-substrate and two-product reaction simplifies to one whose form is similar to that obtained for ordered binding (Eq. 9-44). In the absence of products P and Q Eq. 9-44 will correctly represent the steady-state rate equation corresponding to Eq. 9-50. However, this simplification may not be valid for a very rapidly acting enzyme.

***Isotope exchange at equilibrium.*** Consider the reaction of substrates A and B to form P and Q (Eq. 9-51). If both reactants and both products are present with the enzyme and in the ratio found at equilibrium no net reaction will take place. However, the reactants and products will be continually interconverted under the action of the enzyme. Now if a small amount of

$$A + B \rightleftarrows P + Q \qquad (9\text{-}51)$$

highly labeled reactant (A* or B*) is added, the rate at which isotope is transferred from the labeled reactant into one or the other of the products can be measured. In general, a label in one of the substrates will appear in only one of the products.

Figure 9-7A shows the rate of exchange of isotopically labeled glucose (glucose*) with glucose 6-phosphate as catalyzed by the enzyme hexokinase (Chapter 12). The exchange rate is plotted against the concentration of glucose 6-phosphate with the ratio [glucose] / [glucose 6-phosphate] constant at 1/19, such that an equilibrium ratio for reactants and products is always maintained. As can be seen from the graph, this exchange rate increases monotonically as substrate concentrations are increased. This is also true for the rate of ATP–ADP exchange. The fact that both exchange rates increase continuously indicates random binding of substrates.[55] The inequality of the two maximal exchange rates suggests that release of glucose 6-phosphate may be slower than that of ADP.

Figure 9-7B shows similar plots for lactate dehydrogenase.[53] In this case, after an initial rise (that is not regarded as significant), the pyruvate*–lactate exchange reaches a high constant value as the amount of pyruvate is increased (with a constant [pyruvate] / [lactate] ratio of 1/35). However, the NAD*–NADH exchange increases rapidly at first but then drops abruptly as the pyruvate and lactate concentrations continue to increase. This suggests an ordered mechanism (Eq. 9-43) in which $NAD^+$ and NADH represent A and Q, respectively, and pyruvate and lactate represent B and P. As the concentrations of B and P become very high, the

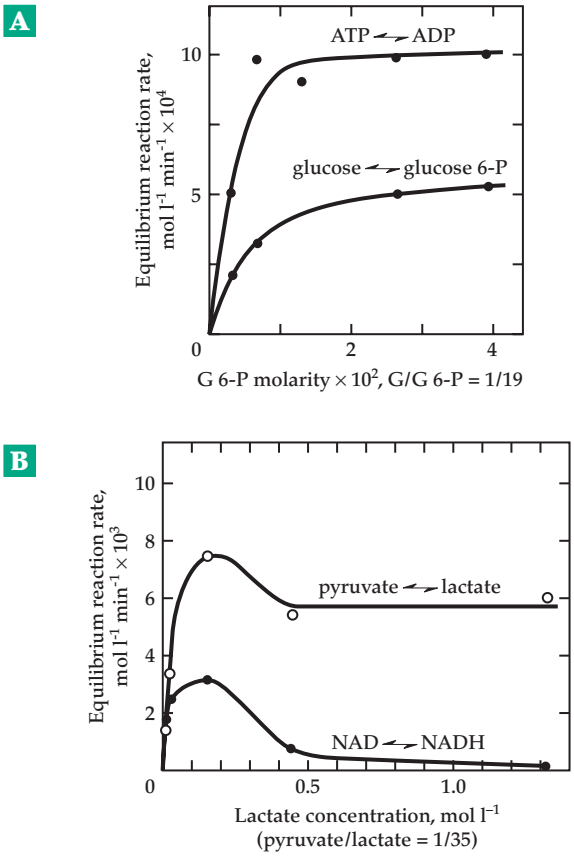



**Figure 9-7** (A) Effect of glucose and glucose 6-phosphate concentrations on reaction rate of yeast hexokinase at equilibrium. Reaction mixtures contain 1–2.2 mM ATP, and 25.6 mM ADP at pH 6.5. From Fromm *et al.*[51] (B) Effect of lactate and pyruvate concentrations on equilibrium reaction rates of rabbit muscle lactate dehydrogenase. Reaction mixtures contained 1.7 mM $NAD^+$, and 30 – 46 µM NADH in Tris-nitrate buffer, pH 7.9, 25°C. From Silverstein and Boyer.[53]

enzyme shuttles back and forth between EA and EQ, but these two complexes rarely dissociate to give free enzyme and A or Q. Hence, the A*–Q exchange rate drops.

In other cases a label may be transferred from A into P or from B into Q. Information on such exchanges has provided a valuable criterion of mechanism which is considered in Chapter 12, Section B,4.

## 5. Kinetics of Rapid Reactions

The fastest steps in an enzymatic process cannot be observed by conventional steady-state kinetic methods because the latter cannot be applied to reactions with half-times of less than about 10 s. Consequently, a variety of methods have been developed[18,56–59a] to measure rates in the range of 1 to $10^{13}$ $s^{-1}$.

***Flowing substrates together.*** One of the first rapid kinetic methods to be devised consists of rapidly mixing two flowing solutions together in a special mixing device and allowing the resulting reaction mixture to move at a rate of several meters per second down a straight tube. At a flow velocity of 10 m $s^{-1}$ a solution will move 1 cm in $10^{-3}$ s. Observations of the mixture are made at a suitable distance, e.g., 1 cm, and with various flow rates. Using spectrophotometry or other observation techniques, the formation or disappearance of a product or reactant can be followed. The special advantage of this technique is that observation can be made slowly. However, it may require large amounts of precious reactant solutions, e.g., those of purified enzymes.

In the **stopped flow** technique two solutions are mixed rapidly by the flow technique during a period of only 1–2 (or a few) milliseconds. A ram drives the solutions from syringes through a mixing chamber into an observation chamber. After the flow stops light absorption, fluorescence, conductivity, or other property, is measured. A means of rapid observation of changes during the reaction is essential. For example, light absorption may be measured by a photomultiplier with data being collected by a computer. Relaxation times as short as a few milliseconds or less can be observed in this way.[59a,b]

***Observing relaxation.*** Kinetic measurements over periods of tens of microseconds or less can be made by rapidly inducing a small displacement from the equilibrium position of a reaction (or series of reactions) and observing the rate of return (relaxation) of the system to equilibrium. Best known is the **temperature jump** method devised by Eigen and associates. Over a period of about $10^{-6}$ s a potential difference of ~100 kV is applied across the experimental solution. A rapid electrical discharge from a bank of condensers passes through the solution (without any sparking) raising the temperature 2–10 degrees. All the chemical equilibria for which $\Delta H \neq 0$ are perturbed. If some property, such as the absorbance at a particular wavelength or the conductivity of the solution, is measured, very small relaxation times can be determined.

While it may not be intuitively obvious, if the displacement from equilibrium is small, the rate of return to equilibrium can always be expressed as a first-order process (e.g., see Eq. 9-13). In the event that there is more than one chemical reaction required to reequilibrate the system, each reaction has its own characteristic relaxation time. If these relaxation times are close together, it is difficult to distinguish them; however, they often differ by an order of magnitude or more. Therefore, two or more relaxation times can often be evaluated for a given solution. In favorable circumstances these relaxation times can be related directly to rate constants for particular steps. For example, Eigen measured the conductivity of water following a temperature jump[18] and observed the rate of combination of $H^+$ and $OH^-$ for which $\tau$ at 23°C equals 37 x $10^{-6}$ s. From this, the rate constant for combination of $OH^-$ and $H^+$ (Eq. 9-52) was calculated as follows (Eq. 9-53):

$$H^+ + OH^- \longrightarrow H_2O \qquad (9\text{-}52)$$

$$k = 1/\{\tau([OH^-] + [H^+])\} = 1.3 \times 10^{11}\ M^{-1}\ s^{-1} \qquad (9\text{-}53)$$

Pressure jump and electric field jump methods have also been used, as have methods depending upon periodic changes in some property. For example, absorption of ultrasonic sound causes a periodic change in the pressure of the system.

***Rapid photometric methods.*** Another useful method has been to discharge a condenser through a flash tube over a period of $10^{-12}$ to $10^{-4}$ s, causing a rapid light absorption in a sample in an adjacent parallel tube. Following the flash, changes in absorption spectrum or fluorescence of the sample can be followed. The availability of intense lasers as light sources has made it possible to follow the results of light flashes of 5–10 picosecond duration and to measure extremely short relaxation times (Chapter 23).[58,59]

***Some results.*** Rapid kinetic methods have revealed that enzymes often combine with substrates extremely quickly,[60] with values of $k_1$ in Eq. 9-14 falling in the range of $10^6$ to $10^8$ $M^{-1}$ $s^{-1}$. Helix–coil transitions of polypeptides have relaxation times of about $10^{-8}$ s, but renaturation of a denatured protein may be much slower. The first detectable structural change in the vitamin A-based chromophore of the light-operated proton pump bacteriorhodopsin occurs in ~5 x $10^{-8}$ s, while a proton is pumped through the membrane in

~$10^{-4}$ s.[61] Interconversion between chair and boat forms of cyclohexane derivatives may have $\tau \sim 10^{-5}$ s at room temperature, while rotation about a C–N bond in an amide linkage may be very slow with $\tau \sim 0.1$ s. The nonenzymatic hydration of the aldehyde pyridoxal phosphate via Eq. 13-1 occurs with $\tau = 0.01 - 0.1$ s, depending upon the pH.[59]

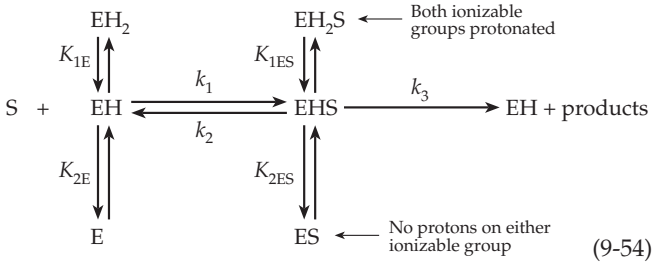
## 6. Cryoenzymology

An alternative to studying rapid reactions is to cool enzymes to a subzero temperature (down to –100°C) where reactions proceed more slowly.[60,62] The enzyme must be dissolved in a suitable "cryosolvent," often containing 50–80% by volume of an organic solvent or solvents such as methanol, ethanol, dimethyl sulfoxide, dimethyl formamide, or ethylene glycol. Those containing methanol are especially desirable because of their low viscosities. Kinetics can be studied by various spectroscopic methods and stopped-flow, temperature-jump, and other rapid-reaction techniques can be applied. One goal of cryoenzymology is to stabilize otherwise unstable intermediates. X-ray crystallographic measurements can also be made at these low temperatures and it should be possible to observe structures of intermediate ES complexes. A problem is that the forms of the complexes stabilized may be side-products rather than true intermediates. However, as discussed in Chapter 3, a combination of low-temperature Laue X-ray diffraction and a laser-induced temperature jump may be feasible.

## 7. The Effect of pH on Enzymatic Action

Because proteins contain many acidic and basic groups it is not surprising that the activity of enzymes often varies strongly with pH. However, it is usually found that the state of protonation of only a few groups has a strong effect on activity. This is understandable because most ionized groups are on the outer surface of protein molecules and most are not close to an active site. Protonation or deprotonation of those groups will hardly ever have a major influence on events in the active site. Often only one or two ionizable groups have highly significant effects.

For many enzymes a plot of $V_{max}$ against pH is a bell-shaped curve (Fig. 9-8). The **optimum** rate is observed at some intermediate pH, which is often, but not always, in the range of pH 6–9. This type of curve can be interpreted most simply by assuming the presence in the active site of two ionizable groups and three forms of the enzyme with different degrees of protonation: E, EH, and $EH_2$.




(9-54)

Let us designate the acid dissociation constants ($K_a$) for the two groups in the enzyme as $K_{1E}$ and $K_{2E}$ and those of the ES complex $K_{1ES}$ and $K_{2ES}$. However, as discussed in Chapter 6, Section E,2 these consecutive $K_a$ values cannot necessarily be assigned to single groups. They may belong to a system of interacting groups. Each may be the sum of more than one microscopic $K_a$ and there may be extensive tautomerism within the active site. The rate constants $k_1$, $k_2$, and $k_3$ define the rates of formation and breakdown of the ES complex.

If it is assumed that only form EHS reacts to give products, the bell-shaped curves of Fig. 9-8 are obtained. The frequent observation of such curves supports the model of Eq. 9-54. It also suggests that the two ionizable groups may be intimately involved in catalysis, one as an acid and the other as a conjugate base (see Section E,5).

For the simple case illustrated in Eq. 9-54, the pH dependence of the initial maximum velocity, the apparent Michaelis constant, and $V_{max}/K_m$ are given by Eqs. 9-55 to 9-57.

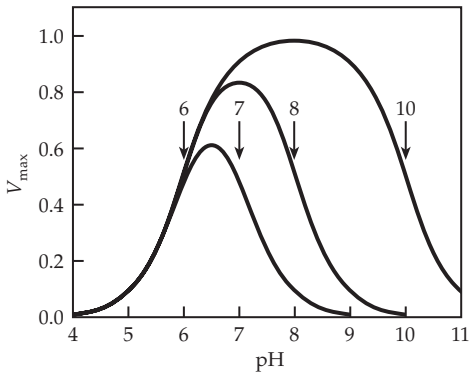



**Figure 9-8** Expected dependence of $V_{max}$ on the pH according to Eq. 9-54 with $k_3[E_t] = 1$, $pK_{1ES} = 6$, and $pK_{2ES}$ as given on the graph. After Alberty.[63] Computer-drawn graph courtesy of Carol M. Metzler.

$$V_{max} = \frac{k_3[E]_t}{(1 + [H^+]/K_{1ES} + K_{2ES}/[H^+])} = \frac{k_3[E]_t}{F_{2ES}} \quad (9\text{-}55)$$

$$K_m = \frac{(k_2 + k_3)}{k_1} \cdot \frac{(1 + [H^+]/K_{1E} + K_{2E}/[H^+])}{(1 + [H^+]/K_{1ES} + K_{2ES}/[H^+])} = K_m F_{2E}/F_{2ES} \quad (9\text{-}56)$$

$$\frac{V_{max}}{K_m} = \frac{k_1 k_3}{(k_2 + k_3)} \cdot \frac{1}{(1 + [H^+]/K_{1E} + K_{2E}/[H^+])} = \frac{k_3[E]_t}{K_m F_{2E}} \quad (9\text{-}57)$$

The denominators (which are the Michaelis pH functions given by the first three terms of Eq. 3-6) represent the fraction of enzyme or of ES complex in the monoprotonated state. The pH dependence of enzymatic action is often more complex than that shown in Fig. 9-8 and given by the foregoing equations. However, it is easy to write Michaelis pH functions (see Chapter 3) for enzymes with any number of dissociable groups in both E and ES and to write appropriate equations of the type of

## BOX 9-B GROWTH RATES OF CELLS

How do we correctly describe the rate at which a cell grows? Consider bacteria in their rapidly growing "log phase."[a] Each cell divides after a fixed length of time, the **doubling time**. For *E. coli* this may be as short as 17 min in the early stages of growth but becomes somewhat longer as time goes on. A mean value of about 26 min. for *E. coli* at 37°C is typical. In contrast, the doubling time for mammalian cells in tissue culture is often about one day.

If a given volume of culture contains $N_0$ cells initially, the number $N_n$ after $n$ cell divisions will be:

$$N_n = 2^n N_0$$

From this equation we can calculate that a single bacterium with a generation time of 20 min can produce $2^{144}$ cells in 48 h of exponential growth. The **exponential growth rate constant *k*** is equal to *the number of doublings per unit time*. Thus, $k$ is the reciprocal of the doubling time. It is easy to show that the number of bacteria present at time $t$ will be given by the following equation.

$$N_t = 2^{kt} N_0$$

This can be rewritten in a form that can be used to determine $k$ by counting the number of bacteria at zero time and at time $t$.

$$kt = \log_2(N_t/N_0) = \log_{10}(N_t/N_0) / 0.301$$

Another way of expressing the growth is to equate the rate of increase of the number of bacteria with a growth rate constant μ multiplied by the number of bacteria present at that time.

$$dN/dt = \mu N$$

This is a general equation for an autocatalytic reaction and $N$ could be replaced with a concentration, for example, the total content of cellular matter per liter in the medium. From the two preceding equations the following can be derived.[a]

$$\mu = k \ln 2 = 0.69k$$

When bacteria are transferred to new medium there is usually a lag before exponential growth begins. Exponential growth eventually stops and the culture enters **stationary phase**, which is usually followed by relatively rapid death of cells in the culture. It is often desirable to study cell growth under conditions of **continuous cultivation** in which a constant generation time is maintained but the density of cells in the medium does not increase. This can be done with a simple device known as the **chemostat**.[a,b] A culture vessel containing bacteria is stirred to ensure homogeneity. Fresh culture medium continuously enters the vessel from a reservoir and part of the content of the vessel, suspended bacteria included, is continuously removed through another tube. The bacterial population in the vessel builds up to a constant level and can be maintained at the same level for relatively long periods of time.

Here are some other statistics on cell growth. Bacteria growing exponentially expand their linear dimensions by 1.5 nm and synthesize $1.6 \times 10^7$ Da of new cell material in one second. This is equivalent to ~1000 small proteins and includes 23 ribosomes and 3000 base pairs of DNA at each growing point. The much larger HeLa human tumor cell grows by 0.13 nm / s but makes $4.6 \times 10^8$ Da / s of material.[c]

---

[a] Stanier, R. Y., Douderoff, M., and Adelberg, E. A. (1970) *The Microbial World*, 3rd ed., Prentice-Hall, Englewood Cliffs, New Jersey (pp. 298–324)

[b] Smith, H. L., and Waltman, P., eds. (1995) *The Theory of the Chemostat; Dynamics of Microbial Competition*, Cambridge University Press, London and New York

[c] Pollard, E. C. (1973) in *Cell Biology in Medicine* (Bittar, E. E., ed), pp. 357–377, Wiley, New York

Eqs. 9-55 to 9-57. Bear in mind that if the **free substrate** contains groups dissociating in the pH range of interest, a Michaelis pH function for the free substrates will also appear in the numerator of Eq. 9-56. If the pH dependence of the enyzme is regulated by a conformational change in the protein, there may be a cooperative gain or loss of more than one proton and the Michaelis pH function must reflect this fact. This can sometimes be accomplished by addition of a term related to Eq. 7-45. For more information see Dixon and Webb,[64] Cleland,[65,66] or Kyte.[67]

Plots of log $V_{max}$ (or log-specific activity) and –log $K_m$ or log ($V_{max}$ / $K_m$) versus pH yield graphs of the type shown in Fig. 9-9. The curved segments of the graphs that extend for about 1.5 pH units on either side of each p$K_a$ are asymptotic to straight lines of slope 1 when a single proton is involved or of a higher slope for multiple cooperative proton dissociation. The straight lines can be extrapolated and intersect at the p$K_a$ values. (However, it is better to fit a complete curve of theoretically correct shape to the points.) Note that the curved line always passes below or above the intersection point at the value of log 2 = 0.30 except in the case of cooperative proton dissociation when it is closer.

Upward turns in the curve of log $K_m$ vs pH correspond to p$K_a$ values in the free enzyme or substrate and downward turns to p$K_a$ values in ES. This approach to analysis of pH dependence has been adopted widely but often incorrectly. For example, many published curves have very sharp bends in which the curved portion covers less than 3 pH units and the curve is much closer than 0.30 to the extrapolated point. This suggests cooperative proton binding and an *apparent* p$K_a$ that is related to $\bar{K}$ of Eq. 7-21. Cooperativity is always a possibility if a conformational change in the protein is involved.[67]

The simple treatment given above is based on the assumption that all proton dissociations are rapid compared to $k_{cat}$, that enzyme in only one state of protonation binds substrate, and that ES in only one state of protonation yields products. These assumptions are not always valid. It also assumes that both binding and dissociation of substrate are rapid, that is, to use Cleland's terminology the substrate is not "sticky." For a sticky substrate that dissociates more slowly than it reacts to form products ($k_3 > k_2$; Eq. 9-54), the values of p$K_{1E}$ will be lowered and p$K_{1E}$ of Eq. 9-53 will be raised by log (1 + $k_3$ / $k_2$).[65,66] In addition to the articles by Cleland, other detailed treatments of pH effects have been prepared by Brocklehurst and Dixon[69] and Tipton and Dixon.[70]

**Fumarate hydratase** (fumarase), which is discussed in Chapter 13, catalyzes the reversible hydration of fumaric acid to malic acid (Eq. 13-11). It was one of the first enzymes whose pH dependence was studied intensively. A bell-shaped pH dependence

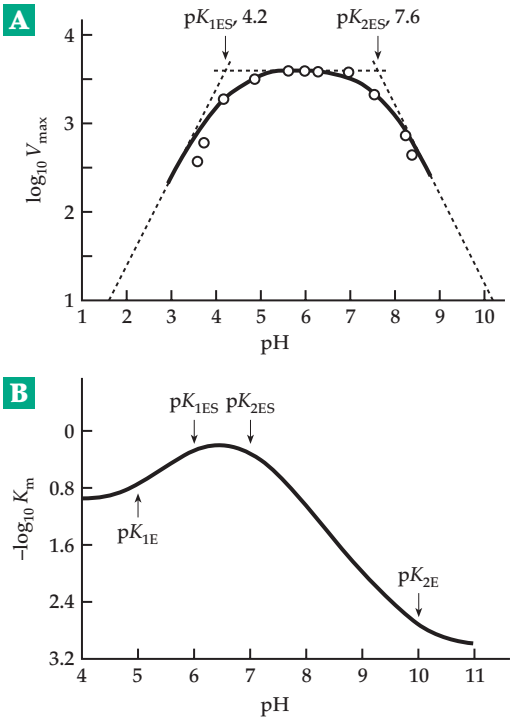



**Figure 9-9** (A) Plot of log $V_{max}$ vs pH for a crystalline bacterial α-amylase. From Ono, et al.[68] (B) Theoretical curve of log $K_m$ vs pH for Eq. 9-56 with p$K_{1E}$ = 5, p$K_{2E}$ = 10, p$K_{1ES}$ = 6, and p$K_{2ES}$ = 7. Courtesy of C. Metzler.

for both forward and reverse reactions was observed by Alberty and coworkers[63] and, using Eqs. 9-54 and 9-55, the two apparent p$K_a$ values were measured.

## B. Inhibition and Activation of Enzymes

The action of most enzymes is inhibited by many substances. Inhibition is often specific, and studies of the relationship between inhibitor structure and activity have been important to the development of our concepts of active sites and of complementarity of surfaces of biomolecules. Inhibition of enzymes is also the basis of the action of a very large fraction of important drugs. Inhibition may be **reversible** or **irreversible**, the latter leading to permanent inactivation of the enzyme. Often, but not always, irreversible inhibition is preceded by reversible binding of the inhibitor at a complementary site on the enzyme surface.

### 1. Competitive Inhibitors

Inhibitors with close structural similarities to a substrate tend to bind to the substrate site. In truly competitive inhibition, substrate and inhibitor not only

compete for the same site but also their binding is reversible and mutually exclusive. The affinity of the inhibitor for the enzyme is expressed quantitatively through the **inhibition constant** $K_i$ which is the *dissociation constant of the enzyme inhibitor complex EI* :

$$K_i = [E][I] / [EI] \tag{9-58}$$

Using the steady state assumption for the mechanism shown in Eq. 9-14, and writing a mass balance equation that includes not only free enzyme and ES but also EI we obtain an equation relating rate to substrate concentration. It is entirely analogous to Eq. 9-15 but $K_m$ is replaced by an apparent Michaelis constant, $K'_m$:

$$K'_m = K_m\left(1 + \frac{[I]}{K_i}\right) \tag{9-59}$$

The relationships between $v$ and [S] and between $1/v$ and $1/[S]$ are as follows:

$$\frac{v}{[S]} = \frac{V_{max}}{K_m'} - \frac{v}{K_m'} \tag{9-60}$$

$$\frac{1}{v} = \frac{1}{V_{max}} + \frac{K_m}{V_{max}[S]}\left(1 + \frac{[I]}{K_i}\right) \tag{9-61}$$

A commonly used test for competitive inhibition is to plot $1/v$ vs $1/[S]$ (Eq. 9-61), both in the absence of inhibitor and in the presence of one or more fixed concentrations of I. The result, in each case, is a family of lines of varying slope (Fig. 9-10) that converge on one of the axes at the value $1/V_{max}$. We see that the maximum velocity is unchanged by the presence of inhibitor. If sufficient substrate is added, the enzyme will be saturated with substrate and the inhibitor cannot bind. The value of $K_i$ can be calculated using Eq. 9-61 from the change in slope caused by addition of inhibitor.

The effect of a fixed concentration of a competitive inhibitor on a plot of $v$ against log [S] (Fig. 9-11) is to shift the curve to the right, i.e., toward higher values of [S], but without any change in shape (or in the value of $V_{max}$).

Another plot, introduced by Dixon,[71] is that of $1/v$ versus [ I ] at two or more fixed substrate concentrations. The student should be able to demonstrate that this plot contains a family of straight lines that intersect at a point to the left of the origin with coordinates $[I] = -K_i$ and $1/v = 1/V_{max}$. This plot may fail to distinguish certain types of inhibition discussed in the next section.[72]

Competitive inhibition is extremely common and has great significance for metabolic control and for the effects of drugs and of poisons. Simple ions are often competitive inhibitors. Since many biochemical substances carry negative charges, anions such as $Cl^-$, $HCO_3^-$, $HPO_4^{2-}$, and acetate$^-$ frequently act as competitive inhibitors.

Two special classes of competitive inhibitors are characterized by **slow binding** and **slow, tight binding** to active sites.[73–77] Among the very tight-binding inhibitors, which may also be slow to dissociate from active sites, are transition state inhibitors discussed in Section D,1.

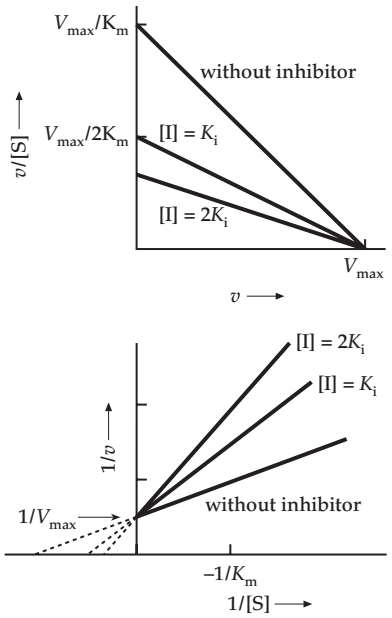



**Figure 9-10** Effect of a competitive inhibitor on the Eadie–Hofstee plot (top) and on a double reciprocal plot (bottom). The apparent $K_m$ (Eq. 9-59) is increased by increasing [I], but $V_{max}$ is unchanged.

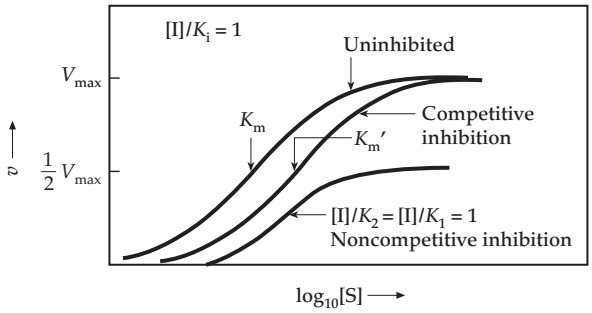



**Figure 9-11** Plots of $v$ vs. log [S] for competitive and noncompetitive inhibition.

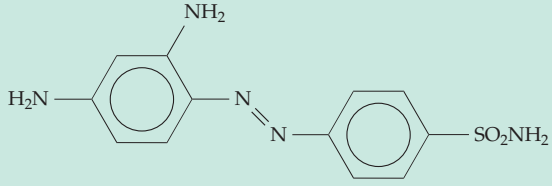
## 2. Noncompetitive Inhibition and Activation; Allosteric Sites

If an inhibitor binds not only to free enzyme but also to the enzyme substrate complex ES, inhibition is **noncompetitive**. In this case, S and I do not mutually exclude each other and both can be bound to the enzyme at the same time. Why does such an inhibitor slow an enzymatic reaction? In most instances, the structure of the inhibitor does not show a close similarity to that of substrate, which suggests that the binding of inhibitors is at an **allosteric site**, that is, at a site other than that of the substrate. The inhibition of the enzyme may result from a distortion of the three-dimensional structure of the enzyme which is caused by the binding of the inhibitor. This distortion may be transmitted to the active site even though the inhibitor binds far from that site. In some cases two distinctly different conformers of the protein may exist, one binding substrate well and the other binding inhibitor well (Fig. 9-12). In other instances the bound inhibitor may interfere with the catalytic action by partially overlapping the active site. In either case the ES complex reacts to give product in a normal way, but the ESI complex reacts more slowly or not at all.

Binding of a substance to an allosteric site sometimes has the effect of *increasing* the activity of an enzyme rather than inhibiting it. This may occur because the **activator** stabilizes the conformation that binds substrate best (Fig. 9-12). The quantitative treatment of such activation is similar to that of inhibition; allosteric inhibitors and activators are often considered together and are referred to as **modifiers** or

## BOX 9-C THE SULFONAMIDES AS ANTIMETABOLITES

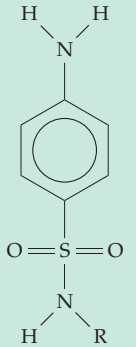
The development of the "sulfa drugs,"[a–c] derivatives of **sulfanilamide**, originated with studies of the staining of protozoal parasites by Paul Erhlich. In 1932 it was shown that the red dye 2,4-diaminoazobenzene-4'-sulfonamide (Prontosil) dramatically cured systemic infections by gram-positive bacteria. Subsequent studies revealed that bacteria converted



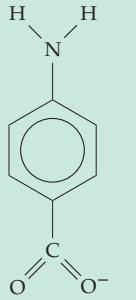
Prontosil

the azo dye to sulfanilamide, a compound with strong bacteriostatic activity; that is, it inhibited bacterial growth without killing the bacteria. Although sulfanilamide had been used in large quantities since 1908 as an intermediate in synthesis of dyes, its potential as an antibacterial agent had not been recognized.

In 1935, D. D. Woods found that the growth inhibition of sulfanilamide was reversed by yeast extract.[d] From this source, in 1940 he isolated ***p*-aminobenzoic acid** and demonstrated that the inhibitory effect of $3 \times 10^{-4}$ M sulfanilamide was overcome by $6 \times 10^{-8}$ M *p*-aminobenzoate. The relationship between the two compounds was shown to be strictly competitive. If the sulfanilamide concentration was doubled, twice as much *p*-aminobenzoate was required to reverse the inhibition as before. These facts led to the formulation by Woods and by P. Fildes (in 1940) of the **antimetabolite theory**.[d,e] It was proposed that *p*-aminobenzoate was needed by bacteria and that sulfanilamide competed for a site on an enzyme designed to act on *p*-aminobenzoate. We now know that the idea was correct and that the enzyme on which the competition occurs catalyzes the synthesis of dihydropteroic acid (Fig. 25-19), a precursor to folic acid.



Sulfonamides
R=H in sulfanilamide



*p*-Aminobenzoate

While sulfanilamide itself is somewhat toxic, a variety of related drugs of outstanding value have been developed. Over 10,000 sulfonamides and related compounds have been tested for antibacterial action.

[a] Bardos, T. J. (1974) *Top. Curr. Chem.* **52**, 63–98
[b] Gale, E. F., Cundliffe, E., Reynolds, P. E., Richmond, M. H., and Waring, M. J. (1972) *The Molecular Basis of Antibiotic Action*, Wiley, New York
[c] Shepherd, R. G. (1970) in *Medicinal Chemistry*, 3rd ed. (Burger, A., ed), pp. 255–304, Wiley (Interscience), New York
[d] Woods, D. D. (1940) *Brit. J. Exp. Pathol.* **21**, 74–90
[e] Fildes, P. (1940) *Lancet* **1**, 955–957

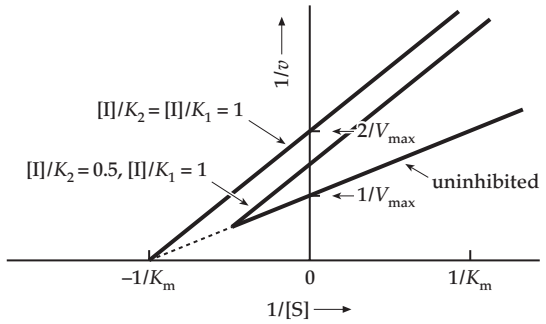
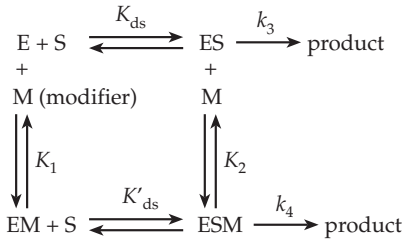



**Figure 9-12** Double reciprocal plots for two cases of noncompetitive inhibition.

**effectors**. A general scheme[78,79] is given by Eq. 9-62.




(9-62)

Here $K_1$ and $K_2$ are equilibrium constants for dissociation of M from EM and ESM, respectively, while $K_{ds}$ and $K'_{ds}$ are the dissociation constant of ES and of ESM (to S and EM), respectively. Notice that, $K'_{ds}$ is not independent of the others (Eq. 9-63):

$$K'_{ds} = K_{ds}K_2/K_1 \tag{9-63}$$

We have already considered competitive inhibition which is obtained when $K_2 = 0$ (and therefore $K'_{ds} = 0$). For this case, M is always an inhibitor and no activation is possible. Notice that the inhibition will appear competitive even if M binds at an allosteric site as in Fig. 9-13 or if the inhibited form does not react at all with substrate. Noncompetitive inhibition will be observed if ESM is formed but does not react, i.e., if $k_4 = 0$. Then the rate equation in reciprocal form will be given by Eq. 9-64.

$$\frac{1}{v} = \frac{1}{V_{max}}\left(1 + \frac{[I]}{K_2}\right) + \frac{K_m}{V_{max}[S]}\left(1 + \frac{[I]}{K_1}\right) \tag{9-64}$$

We see that $1/V_{max}$ is multiplied by a term containing [I] and $K_2$. It is a characteristic of noncompetitive inhibition that the maximum velocity is decreased from that observed in the absence of the inhibitor. Also, no matter how high the substrate concentration, inhibition cannot be reversed completely.

Figure 9-12 shows a plot of $1/v$ against $1/[S]$ at a series of fixed values of [I]. For the case that $K_1 = K_2$ **(classical noncompetitive inhibition)**, a family of reciprocal plots that intersect on the horizontal axis at a value of $-1/K_m$ is obtained. On the other hand, if $K_1$ and $K_2$ differ (the general case of noncompetitive inhibition), the family of curves intersect at some other point to the left of the vertical axis and, depending upon the relative values of $K_1$ and $K_2$, either above or below the horizontal axis. The example illustrated is for $K_2 = 0.5K_1$; that is, for the binding of M to ES being twice as strong as that to E.

Figure 9-11 shows inhibition data for both the noncompetitive and the competitive cases plotted vs log[S]. The shift of the midpoint to the right in each case reflects the tendency of the inhibitor to exclude the substrate from binding, while the lowered value of the maximum velocity in the case of noncompetitive inhibition results from the failure of the substrate to completely displace the inhibitor from the enzyme

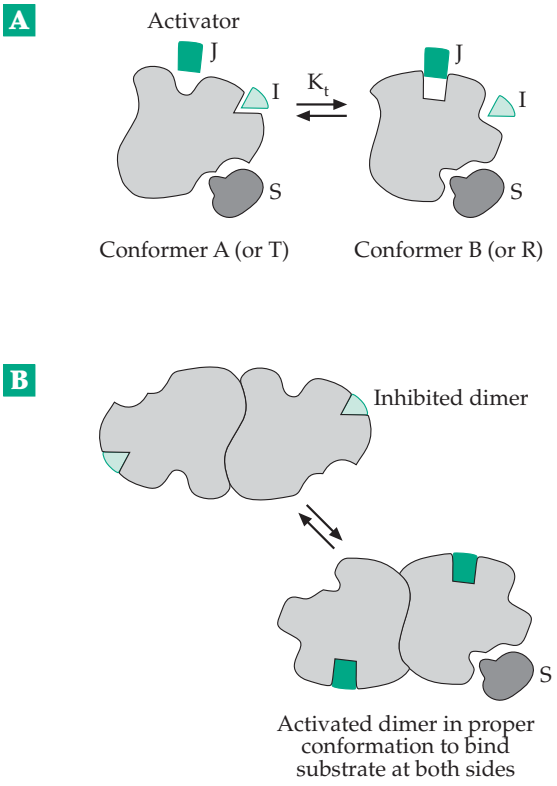



**Figure 9-13** (A) An enzyme with binding sites for allosteric inhibitor I and activator J. Conformer A binds inhibitor I strongly but has little affinity for activator J or for substrate S. Conformer B binds S and catalyzes its reaction. It also binds activator J whose presence tends to lock the enzyme in the "on" conformation B. Conformers A and B are designated T and R in the MWC model of Monod, Wyman, and Changeux.[80] (B) Inhibited and activated dimeric enzymes.

even as [S] becomes very high. If an inhibitor binds only to ES and not to E, i.e., $K_1 = 0$, a family of parallel double-reciprocal plots of $1/v$ vs $1/[S]$ will be obtained. This case is referred to as **uncompetitive inhibition**. Multisubstrate enzymes with either ordered sequential or ping-pong mechanisms also often give parallel line plots with inhibitors.

Uncompetitive inhibitors of liver alcohol dehydrogenase (Chapter 15) could be used to treat cases of poisoning by methanol or ethylene glycol.[81–83] The aim is to prevent rapid oxidation to the toxic acids HCOOH and $HOCH_2COOH$, which lower blood pH, while the alcohols are excreted. Uncompetitive inhibitors have an advantage over competitive inhibitors as therapeutic agents in that the inhibition is not overcome when the substrate concentration is saturating.[84]

If rate constant $k_4$ of Eq. 9-62 is not zero we may observe either inhibition or activation. If $k_3 = k_4$ the effect of a modifier will be to alter the apparent $K_m$ by either increasing it (inhibition) or decreasing it (activation). The maximum velocity will be unchanged. Monod *et al*.[80] referred to enzymes showing such behavior as **K systems**. On the other hand, if $K_1 = K_2$ and $k_3$ differs from $k_4$ we have a purely **V system**. In the general case a modifier affects both the apparent $K_m$ and $V_{max}$. Although in the foregoing discussion it has been assumed that activators bind at allosteric sites, as was pointed out in Section A,11, ions and other small molecules (substrate surrogates) that act as competitive inhibitors at high concentrations may be activators at low concentrations.[40]

Activation of enzymes by specific metallic ions is often observed. In many instances the metallic ion is properly regarded as a *second substrate* which must bind along with the first substrate before reaction can occur. Alternatively, the complex of the organic substrate with the metal ion can be considered the "true substrate." Thus, many enzymes act upon the magnesium complex of ATP (Chapter 12). The enzymes can either be regarded as three-substrate enzymes requiring $Mg^{2+} + ATP^{4-}$ and another substrate or as two-substrate enzymes acting upon $ATPMg^{2-}$ and a second substrate.

## 3. Inhibitors in the Study of Mechanisms

A substrate analog will frequently inhibit only one of the two forms of a multisubstrate enzyme with a ping-pong mechanism.[1,72] Reciprocal plots made for various inhibitor concentrations consist of a family of parallel lines reminiscent of uncompetitive inhibition. Observation of such parallel line plots can support a ping-pong mechanism for an enzyme but cannot prove it because in some cases parallel lines are observed for inhibition of enzymes acting by an ordered sequential mechanism. The following question arises naturally for any ordered bimolecular reaction (Eq. 9-43): Of the two substrates required by the enzymes, which one binds to the enzyme first? If the concentration of one substrate is kept constant while varying concentrations of an inhibitory analog of that substrate are added and $1/v$ is plotted against the reciprocal of the concentration of the other substrate, parallel lines are obtained if, and only if, the substrate of fixed concentration is B, the substrate is adding second in the binding sequence, *and* if I is its analog. The substrate binding first (A) is the one whose concentration was varied in the experiment.

Product inhibition (Section A,12) can also provide information about mechanisms. For example, if $1/v$ is plotted against $1/[A]$ in the presence and absence of the product Q, the product will be found to compete with A and to give a typical family of lines for competitive inhibition. On the other hand, a plot of $1/v$ vs $1/[B]$ in the presence and absence of Q will indicate noncompetitive inhibition if the binding of substrates is ordered (Eq. 9-43). In other words, only the A–Q pair of substrates are competitive. Product inhibition is also observed with enzymes having ping-pong kinetics (Eq. 9-47) as a result of formation of nonproductive complexes.

## 4. Allosteric Effectors in the Regulation of Enzyme Activity

The binding of a substance at an allosteric site with the induction of a conformational change forms the basis for many aspects of regulation. The term **allostery** (allosterism) usually refers to the effects of allosteric modifiers, which may be either inhibitors or activators, on oligomeric enzymes. However, as we have already seen (Eq. 9-62), monomeric enzymes may also be subject to allosteric regulation by modifiers. Consider a monomer that contains binding sites for substrate, inhibitor, and activator and which exists in conformations A and B as in Fig. 9-13. Let us assume (see Eq. 7-31) that conformer B binds both substrate and activator well but that it binds inhibitor poorly or not at all. On the other hand, A binds inhibitor well but binds substrate and activator poorly. This simple combination of two conformers with different binding properties provides a means by which enzymes can be turned "on" or "off" in response to changing conditions.

If an inhibitory substance builds up to a high concentration within a cell, it binds to conformer A; if the inhibitor concentration is high enough, virtually all of the enzyme will be locked in the inactive conformation A. The enzyme will be turned off or at least reduced to a low activity. On the other hand, in the presence of a high concentration of activator the enzyme will be turned on because it is locked in the B conformation. The relative concentrations of inhibitor, activator, and substrate within a cell at any given time

will determine what fraction of the enzyme is in active conformation B. It is this interplay of inhibitory and activating effects that provides the basis for much of the regulation of cell chemistry.[80,85]

The effects of inhibitors or activators on the kinetics of the monomeric enzyme of Fig. 9-13 can be described by Eq. 9-62 to 9-64. Separate terms for both inhibition and activation can be included. The equilibrium between the two conformers can also be indicated explicitly according to Eq. 7-30. However, for monomeric enzymes it is usually not profitable to try to separate the two constants $K_t$ and $K_{BX}$ which describe the conformational change and binding of substrate or activator, respectively, in Eq. 7-30.

Most intracellular enzymes are oligomeric, and the binding of allosteric effectors leads to additional interesting effects. Binding constants or dissociation constants must be defined for both inhibitor and activator to both conformers A and B. Since all species must be taken into account in the mass balance, the equations are complex. However, the Monod–Wyman–Changeux (MWC) model (Chapter 7) gives a relatively simple picture. The saturation curve for an oligomeric enzyme following this model may be derived from Eq. 7-39 and is given by the following expression:

$$Y = \frac{Lc\alpha(1 + c\alpha)^{n-1} + \alpha(1 + \alpha)^{n-1}}{L(1 + c\alpha)^n + (1 + \alpha)^n} \qquad (9\text{-}65)$$

Here, $L$ is the **allosteric constant** which is given (for a dimer) by Eq. 7-36. The constant $c$ is the ratio of dissociation constants $K_{BS}$ and $K_{AS}$ for the two conformers:

$$c = K_{BS} / K_{AS} \qquad (9\text{-}66)$$

Notice that in this chapter *dissociation* constants ($K_d$) of ES complexes are being used, whereas the equations of Chapter 7 are all written in terms of association constant ($K_f$). The parameter α is defined as follows, where $K_{BS}$ equals $1/K_{BX}$ of Chapter 7.

$$\alpha = [S] / K_{BS}) \qquad (9\text{-}67)$$

To take account of effects of inhibitor and activator the *ratios* of dissociation constants of I from BI and AI and of activator J from BJ and AJ are defined as in Eq. 9-68. Likewise, "normalized concentrations" of I and J are defined (Eq. 9-69) as β and γ, respectively.

$$d = \frac{K_{BI}}{K_{AI}} > 1 \qquad e = \frac{K_{BJ}}{K_{AJ}} < 1 \qquad (9\text{-}68)$$

$$\beta = [I]/K_{AI} \qquad \gamma = [J]K_{BJ} \qquad (9\text{-}69)$$

According to the MWC model, in the presence of inhibitor and activator at normalized concentrations β and γ an enzyme will still follow Eq. 9-65, but the allosteric constant $L$ will be replaced by an apparent allosteric constant $L'$ (Eq. 9-70).[86] Figure 9-14 shows plots of Y vs. log α for two different values of $L'$ for a tetramer with a specific value assumed for $c$. In both

$$L' = L\left[\frac{(1 + \beta d)(1 + \gamma e)}{(1 + \beta)(1 + \gamma)}\right]^n \qquad (9\text{-}70)$$

cases, Y approaches 1 as log α increases, but since we are dealing with noncompetitive inhibition at high values of $L'$, much of the enzyme will be in the T(A) conformation at saturation.

Noncompetitive inhibition cannot be completely reversed by very high substrate concentrations. Monod *et al*. defined for an allosteric enzyme a **function of state $\overline{R}$** (Eq. 9-71) which is the fraction of total enzyme in the $\overline{R}$ (B) conformation:

$$\overline{R} = \frac{(1 + \alpha)^n}{L(1 + c\alpha)^n + (1 + \alpha)^n} = \text{function of state} \qquad (9\text{-}71)$$

In a $K$ system (Section B,2) it is the value of $\overline{R}$ that determines the velocity with which an enzyme reacts. Figure 9-14 also shows $\overline{R}$ as a function of log [α]. Note that when $L'$ is low $\overline{R}$ does not approach zero even when [S] → 0. In other words, the enzyme is never completely turned off, just as when $L'$ is high the enzyme is never completely turned on.

Figure 9-14 may be compared with Fig. 9-11 which shows similar curves for noncompetitive inhibition of a monomeric enzyme. The significant difference between the two figures is that saturation of the oligomeric enzyme occurs over a narrower concentration range than does that of the monomer, i.e., saturation of the oligomeric enzyme, especially in the presence of inhibitor, is *cooperative*. Note that cooperative binding of substrate requires that the free enzyme be largely in conformation $T$ (A), as it is in the presence of an inhibitor. Allosteric interactions between two identical molecules, whether of substrate or of effector, are described by Monod *et al*.[80] as **homotropic interactions**. Such interactions lead to cooperativity or anticooperativity in binding. Allosteric interactions between two different molecules, e.g., a substrate and an activator are designated **heterotropic**.

For many enzymes the MWC model is unrealistically simple. The more general treatment of binding equilibria given in Chapter 7 may be applicable. However, in addition to $K$ systems there are $V$ systems in which a conformational change alters the maximum velocity (see Eq. 9-62)[87] and sometimes both substrate

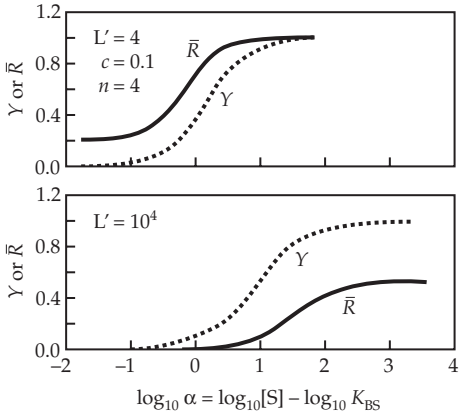
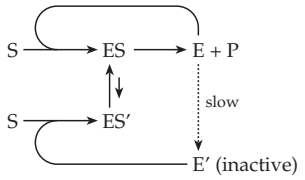



**Figure 9-14** Fractional saturation Y and "function of state" $\bar{R}$ for hypothetical tetrameric enzymes following the MWC model. Curves are calculated for two different values of the apparent allosteric constant $L'$ (Eq. 9-70) and for $c = 0.1$ (Eq. 9-66). After Rubin and Changeux.[86]

affinity and maximum velocity.

The fact that data can be fitted to an equation is not proof that a mathematical model is correct; other models may predict the results just as well. For example, Jenkins has shown that the presence of ions that act as substrate surrogates (Section A,11) can produce cooperative or anticooperative binding curves for substrates.[40,88] Rabin[89] suggested the following explanation of cooperativity for a monomomeric enzyme with a single substrate-binding site. In its active conformation E the enzyme reacts with substrate rapidly and the ES complex yields product rapidly as in the upper loop of Eq. 9-72. However, a slow conformational change interconverts E and E′, a less active form with much lower affinity for substrate.




(9-72)

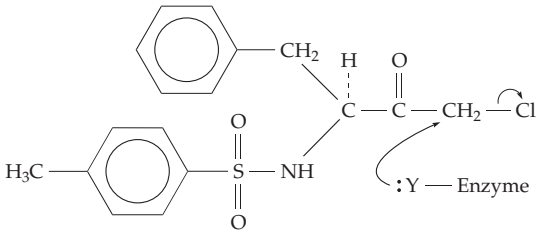
At the same time, the complex ES′, if formed, can equilibriate with ES and thereby alter the conformational state of the protein. At low substrate concentrations E′ will predominate and the enzymatic activity will be low. At high substrate concentrations most E will exist as ES and after release of product will tend to remain in the active conformation long enough to bind and act on another substrate molecule. This and other kinetic models of cooperativity are discussed by Newsholme and Start.[90]

The physiological significance of cooperative binding of substrates to enzymes is analogous to that of cooperative binding of oxygen by hemoglobin which provides for more efficient release of oxygen to tissues. However, in the presence of excess activator an enzyme is locked in the *R* (B) conformation and no cooperativity is seen in the binding of substrate. In this case, each binding site behaves independently. On the other hand, *there will be strong cooperativity in the binding of the activator*. The result is that control of the enzyme is sensitive to a higher power than the first of the activator concentration. Likewise, the turning off of the enzyme is more sensitive to inhibitor concentration as a result of cooperative binding of the latter. It seems likely that the evolution of oligomeric enzymes is at least partly a result of the greater efficiency of control mechanisms based on cooperative binding of effectors.

## 5. Irreversible Inhibition of Enzymes

Some of our most effective drugs, such as penicillin (Box 20-G) are irreversible inhibitors of specific enzymes. Such inhibitors are also of practical importance to biochemists who wish to inhibit specific enzymes such as proteases that might otherwise destroy proteins they are studying. Transition state analogs (Section E,1) and slow-binding inhibitors (Section B,1) may appear to inactivate an enzyme irreversibly because they bind tightly and dissociate slowly. However, true irreversible inhibitors, such as oxidizing agents or alkylating agents, cause a more permanent chemical modification, usually at or near the active site. Two groups of irreversible inhibitors are of special interest: affinity labeling agents and enzyme-activated inhibitors.

**Affinity labeling agents** (active-site directed inactivating reagents) have two essential properties: *a high affinity for the active site* of the specific enzyme to be inhibited and *the presence of a chemically highly reactive group* which can attack a functional group in a protein. A good example is provided by derivatives of chloromethylketones, which are reactive alkylating agents. One of these, *N*-tosylphenylalanine chloromethyl ketone (TPCK), is a potent inhibitor of chymotrypsin. In addition to the chloromethyl ketone group it contains



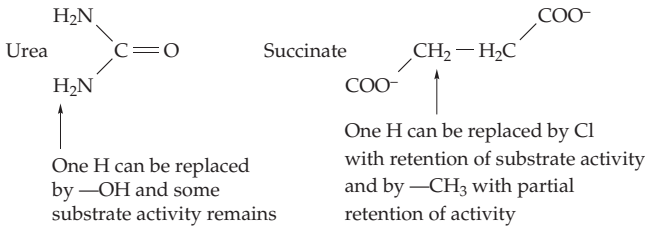

L-*N*-tosylphenylalanine chloromethylketone (TPCK)

the phenylalanine side chain, which helps direct it into the substrate binding site of chymotrypsin, and a sulfonamide linkage that mimics the normal amide linkages of the substrate. The group Y, which becomes alkylated in an $S_N2$-like displacement of a chloride ion (see Eq. 9-76) is thought to be His 57 of the active site (see Fig. 12-10). For additional discussion see Walsh,[91] Kyte,[67] and Plapp.[91a]

**Enzyme-activated inhibitors** (also called suicide substrates, $k_{cat}$ inhibitors, or mechanism-based inhibitors) are chemically inert until they enter the active site of their target enzymes. Then, by passing through at least some of the normal stages of the catalytic action of the enzyme, they are converted to reactive intermediates that can become irreversibly bound to an enzyme.[92,93] For example, a halogen atom (F, Cl, Br) together with a proton may be eliminated from an intermediate to give an unsaturated compound to which a nucleophilic side chain of the protein adds. Several of these inhibitors are discussed in later chapters. Because of their high specificity many enzyme-activated inhibitors are potential drugs. One of them, α-difluoromethylornithine (Box 14-C), is said to bind covalently to only one protein in the body of a rat, namely, ornithine decarboxylase, the target enzyme for the drug.

## C. The Specificity of Enzymatic Action

Enzymes are usually impressively specific in their action. The specificity toward substrate is sometimes almost absolute. For many years urea was believed to be the *only* substrate for the enzyme **urease** and succinate the only substrate for **succinate dehydrogenase**. Even after much searching for other substrates, only




one or two closely related compounds could be found that were acted on at all. In other cases enzymes can use a class of compounds as substrates. For example, the **D-amino acid oxidase** of kidney oxidizes a variety of D-amino acids but does not touch L-amino acids.

Almost as impressive as the substrate specificity of enzymes is the specificity for a given type of reaction. Many substrates are capable of undergoing a variety of different chemical reactions, either unimolecular or with water or some other compound present in the cell. The enzyme catalyzes only one of these reactions. Although side reactions may occur to a small extent, the most impressive thing in comparing an enzyme-catalyzed reaction with an uncatalyzed organic reaction is that the latter often produces large amounts of side reaction products, but the enzymatic reaction does not.

### 1. Complementarity of Substrate and Enzyme Surfaces

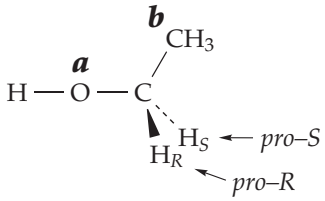
Impressed by the specificity of enzymatic action, biochemists early adopted a "lock-and-key" theory which stated that for a reaction to occur the substrate must fit into an active site precisely. Modern experiments have amply verified the idea. A vast amount of kinetic data on families of substrates and related competitive inhibitors support the idea and numerous X-ray structures of enzymes with bound inhibitors or with very slow substrates have given visual evidence of the reality of the lock-and-key concept. Directed mutation of genes of many enzymes of known three-dimensional structure has provided additional proof.

As anticipated, hydrophobic groups of substrates or inhibitors usually contact hydrophobic regions of the protein and the fit is tight. For example, the active site of chymotrypsin contains a "hydrophobic pocket" designed to hold a large hydrophobic side chain, thus providing the specificity observed with this enzyme (Table 3-2). Likewise, polar groups of substrates contact polar groups of the enzyme. The interactions are complementary, positive charges fitting against negative and with correctly formed hydrogen bonds. Trypsin, whose structure is similar to that of chymotrypsin, exerts its specificity for a positively charged side chain next to the bond cleaved (Table 3-2) by virtue of the presence of a negatively charged carboxylate at the bottom of the hydrophobic pocket. Several C=O and N–H groups of the peptide linkages in the substrate form hydrogen bonds to the edge of a β sheet in the protein, in effect making the substrate an added β strand in the sheet (Fig. 12-10). Aspartate aminotransferase acts on the dicarboxylic amino acids glutamate and aspartate. A pair of arginine side chains bind the two carboxylates of the substrate while the $-NH_3^+$ of the substrate is attracted to a negatively charged group in the coenzyme pyridoxal phosphate, which is also present in the active site (Eq. 14-39, Fig. 14-10).

### 2. Stereospecificity and Prochiral Centers

Most enzymes possess an infallible ability to recognize the difference between the right side and the left side of an organic substrate even when the latter has perfect bilateral symmetry. In fact, this ability is limited to **prochiral centers** of molecules and is a natural consequence of their reaction with the chiral

enzyme.[94–98] Consider carbon atom number 1 of ethanol. The two attached hydrogen atoms are chemically identical and would react identically with a nonchiral reagent. Nevertheless, these atoms are no more equivalent stereochemically than are your right and your left arms. We say that the molecule has a prochiral center at C-1. A prochiral center on a tetrahedrally bonded carbon atom always contains two identical atoms or groups but a total of three different kinds of atoms or groups. If the priority of this H is elevated, e.g., by substitution of $^2H$ for $^1H$ the configuration would be *R* (the priority of this $^2H$ would become **c**).



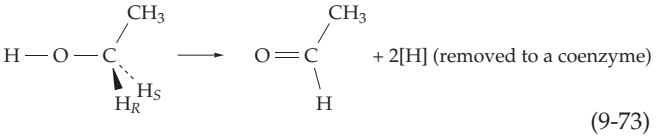

## BOX 9-D RECEPTORS, AGONISTS, AND ANTAGONISTS

Inhibition of enzymes provides the basis for many of the effects of antibiotics and other chemotherapeutic substances (e.g., see Box 9-B). However, some drugs act on cell surface **receptors** which have ordinarily not been regarded as enzymes. According to receptor theory, developed around 1937, structurally similar drugs often elicit similar responses because they bind to the same receptor. A receptor may normally bind a hormone, neurotransmitter, or other metabolite whose geometry is partially shared by a drug. Binding of drugs of one class, termed **agonists** in the pharmacological literature, to an appropriate receptor triggers a response in a cell, similar to that of a hormone. On the other hand, compounds of related structure often act as **antagonists**, binding to receptor but failing to elicit a response. Agonist and antagonist often act in a strictly competitive fashion as in competitive inhibition of enzyme action.

We know now that many receptors *are* enzymes, some of which may act quite slowly. The active site may sometimes be far from the receptor binding site and sometimes in a separate subunit. The receptor can be viewed as an allosteric effector which binds at a distant site or as a ligand for a regulatory subunit of the enzyme complex. Alternatively, the active site may be viewed as the site for relaying a signal received from the hormone or other agonist.

The two hydrogen atoms on C-1 of ethanol are called **enantiotopic** because replacement of one or the other by a fourth different kind of atom or group would produce a pair of enantiomers (page 41). This fact suggests a way of naming the positions occupied by two enantiotopic atoms or groups. We first assign priorities to all of the atoms or groups attached to the central carbon atom according to the *RS* system. Now, we ask whether the configuration will be *R* or *S* when the priority of one of the two identical groups is raised, e.g., by substitution of one of the hydrogen atoms by deuterium. If the configuration becomes *R*, that group occupies a *pro-R* position; if the configuration becomes *S*, it occupies a *pro-S* position. Referring to the preceding diagram, it is easy to see (by viewing down the bond to the group of lowest priority and applying the usual rule for determining configuration) that if the *pro-R* hydrogen ($H_R$) is replaced by deuterium ($^2H$), the configuration will be *R*. Conversely, replacement of $H_S$ by deuterium will lead to the *S* configuration.

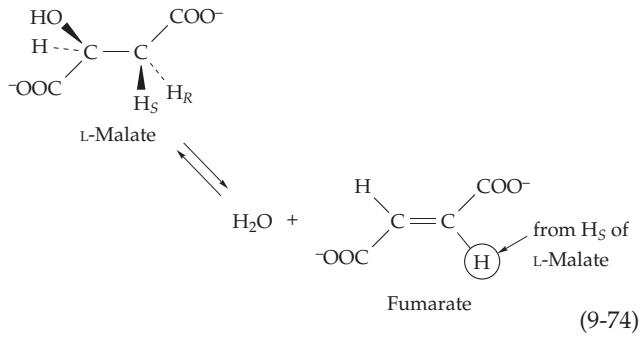
When ethanol is oxidized by the action of **alcohol dehydrogenase** (Eq. 9-73), only the *pro-R* hydrogen atom is removed. If the reaction is reversed in such a way that deuterium is introduced into ethanol from the reduced coenzyme the optically active *R*-2-deuterioethanol is formed. The ability of an enzyme to




(9-73)

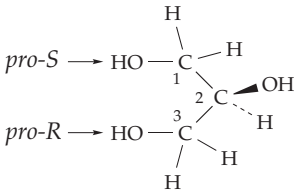
recognize a single hydrogen of a pair of hydrogens on a $CH_2$ group was at first a surprise to many biochemists.[99] However, it is a natural result of the complementarity of enzyme and substrate surfaces, just as the fit of a shoe is determined by the complementarity of surfaces of foot and shoe. Only a chiral catalyst can have this ability.

Another example is provided by malic acid, a chiral molecule which also contains a prochiral center (see Eq. 9-74). In this case replacement of the *pro-R* or *pro-S* hydrogen atom by another atom or group would yield a pair of diastereoisomers rather than enantiomers. Therefore, these hydrogen atoms are **diastereotopic**. When L-malic acid is dehydrated by fumarate hydratase (Chapter 13) the hydrogen in the *pro-R* position is removed but that in the *pro-S* position is not touched. This can be demonstrated by allowing the dehydration product, fumarate, to be hydrated to malate in $^2H_2O$ (Eq. 9-74). The malate formed contains deuterium in the *pro-R* position. If this malate is now isolated and placed with another portion of enzyme in $H_2O$, the deuterium is removed cleanly. The fumarate produced contains no deuterium.




(9-74)

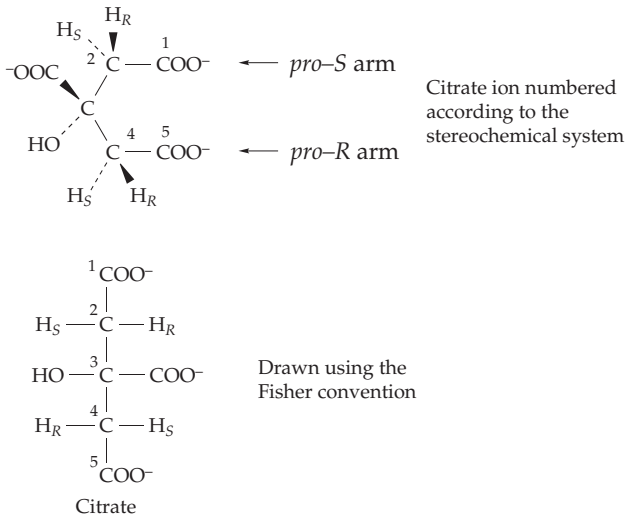
***Stereochemical numbering.*** In some prochiral molecules, such as glycerol, the two ends of the main carbon chain form identical groups. Since the two ends are distinguishable to an enzyme, it is important to decide which should be labeled C-1 and which C-3. Hirschmann proposed a stereochemical numbering system[100] in which the carbons are numbered beginning with the end of the chain that occupies the *pro-S* position.



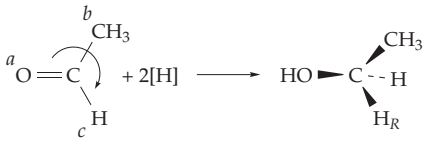

Stereochemical numbering for a symmetric prochiral molecule. Carbons are numbered beginning with the end of the chain that occupies the *pro-S* position.

In this numbering system derivatives of the parent prochiral compound are given the prefix *sn*-. Thus, glycerol phosphate, used by cells to construct phospholipids, usually bears a phosphate group on the $-CH_2OH$ in the *pro-R* position of glycerol and is therefore *sn*-3-glycerol phosphate.

The **citrate ion**, a very important prochiral metabolic intermediate, has three prochiral centers at C-2, C-3, and C-4, respectively. That at C-3 distinguishes the *pro-R* and *pro-S* arms and determines the stereochemical numbering. Citrate containing $^{14}C$ in the *sn*-1 position is called *sn*-citrate[1-$^{14}$C] and is the form of labeled citrate that is synthesized in living cells from oxaloacetate and [1-$^{14}$C]acetyl coenzyme A (see Fig. 10-6). The first step in the further metabolism of citrate is the elimination of the –OH group from C-3 together with the $H_R$ proton from C-4 through the action of the enzyme **aconitate hydratase** (aconitase). In this case the proton at C-4 (in the *pro-R* arm) is selected rather than that at C-2.




***Trigonal prochiral centers.*** Planar trigonal atoms, such as those of aldehydes, ketones, and alkenes, are also prochiral if they are attached to three different kinds of atoms or groups. The two faces are enantiotopic or diastereotopic, if another chiral center is present. Hanson[94,94a] proposed that the faces be named as indicated in Eq. 9-75. The trigonal atom is viewed from one side and the three groups surrounding the carbon atom are given priorities, *a*, *b*, and *c*, as in the *RS* system (Chapter 2). If the sequence *a*, *b*, and *c* of priorites is clockwise, the face toward the reader is *re* (*rectus*); if counterclockwise, it is *si* (*sinister*). Priorities can often be assigned on the basis of the atomic numbers of the first atoms in the three groups (as in the example shown)

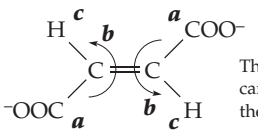



(9-75)

but, if necessary, replica atoms must be added as described in Chapter 2. Replica atoms attached to an originally trigonal carbon are ignored completely, but those attached to O or N of C=O or C=N may be required to establish the priorities of the groups around the carbon atom. For C=O and C=N the faces of the O and N atoms are taken to be the same as that of the attached trigonal carbon atom. In C = C the two carbons may have their *re* faces either on the same side or on opposite sides of the group.

Notice that addition of an atom of $^2H$ to the *re* face of acetaldehyde would give *R*-deuteroethanol (Eq. 9-75; reverse of Eq. 9-73). This is the reaction catalyzed by alcohol dehydrogenase. Addition of $^1H$ to the *re* face places the entering hydrogen in the *pro-R* position. Addition to the *si* face would place it in the *pro-S*

position. Fumarate has two trigonal carbon atoms. The *si* faces of both are toward the viewer in the following structure (also shown in Eq. 9-74). Referring to Eq. 9-74, we see that an $HO^-$ ion from water adds to the *si* face of one carbon atom of fumarate to give *s*-malate (L-malate). At the same time, a proton combines with the *re* face of the adjacent unsaturated carbon atom to enter the *pro-R* position.




The *si* faces of both trigonal carbon atoms are toward the viewer.

Fumarate

## 3. Induced Fit and Conformational Changes

Results of many X-ray studies indicate that the lock-and-key picture of enzyme action must be modified. If an enzyme is a "lock" and the substrate the "key," *the entrance of the key into the lock often induces a conformational change in the protein*. Binding of substrates may be imperfect in the initially formed ES complex but may be more nearly perfect a few nanoseconds or microseconds later as the protein readjusts its structure to accommodate the substrate. The substrate has induced a fit. For example, when an amino acid substrate binds to aspartate aminotransferase one whole domain of the enzyme moves inward, packing hydrophobic side chains of the protein against the substrate (Chapter 14). This strengthens the electrostatic interactions between the ion pairs that orient the substrate and align it for reaction. Similar conformational changes have been observed for citrate synthase, glycogen phosphorylase, various kinases (Chapter 12), alcohol dehydrogenase (Chapter 15), and a growing list of other enzymes. Accompanying changes in circular dichroism, ultraviolet spectra, and sedimentation constants are often observed.

For many other enzymes the observed conformational changes are subtle. A single loop of polypeptide chain or even of a side chain moves to cover the bound substrate. With trypsin and other serine proteases a flexible segment of the peptide chain becomes immobilized and forms more hydrogen bonds after substrate binds than before. It is probably quite unusual for a major unfolding and refolding of parts of a protein to take place. The term induced fit usually refers to substrate binding, but as substrates are interconverted within active sites successive changes in geometry and in charge distribution occur. Small conformational changes may be required at several stages of an enzymatic reaction to ensure that complementarity of substrate and enzyme is preserved.[101] In line with this idea are the facts that proteins are less tightly packed at active sites than in other parts of the molecules, and that active sites often lie between domains and usually are formed by several loops of the peptide chain (Chapter 2).[102]

Although conformational changes allow proteins to maintain good complementarity with substrates, it does not follow that substrates are therefore bound very tightly. This is easiest to understand for reversible reactions in which a substrate is the product of the reverse reaction. Not only binding but also the rate of release of product must be rapid. Very tight binding would retard release and cause product inhibition. Furthermore, as we can see from Eq. 9-37, the velocity of catalysis is the product of [S][E] x $k_{cat} / K_m$, where [E] is, *free enzyme*. Catalysis can be made more efficient in two ways: by increasing $k_{cat}$ or by decreasing $K_m$, that is, by tighter binding. However, tighter binding will be advantageous only if $K_m$ is not so low that the enzyme is approaching saturation with substrate. Otherwise, [E] will fall as $K_m$ is lowered and little increase in rate will be observed.[2]

In fact, it appears that enzymes have evolved to have high values of $k_{cat}$ and *high* values of $K_m$, that is, *weak binding* of substrates.[2] Frequently, in one conformational state of a protein the active site is open, with solvent molecules or substrate surrogates present, while in a second state of nearly equal energy it is closed around the substrate. The functional groups of the enzyme's active site can be bound either to external ions or to ionic groups of substrates and either to water or to hydrogen-bonding groups of substrates. In some cases two ionic groups in the protein may pair with each other in the open conformation and with ionic groups of the substrate in the closed conformation. Thus, the energy changes accompanying the conformational changes can be small but very good complementarity can exist in the ES complex, an important factor in establishing specificity.

## 4. Specificity and $k_{cat}$

Enzyme specificity is often observed not only in binding but also in the rate at which ES is converted to products. Thus, it is the values of $k_{cat} / K_m$ that determine specificity. Good examples are provided by chymotrypsin and related serine proteases (Chapter 12),[2] for which substrates with the shortest chains are often bound as well as those with longer chains but react more slowly. For example, *N*-acetylphenylalanine amide binds to chymotrypsin about as well as does the longer *N*-acetylphenylalanylalanine amide but reacts only 1/47 as fast.[2] One might anticipate that increasing the length of the substrate would make it bind more tightly because of the greater number of contacts between substrate and enzyme. It has often been suggested that the reason that this does *not* happen is that the

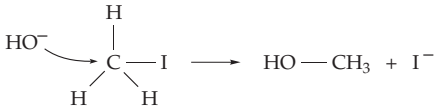
binding of the longer, more specific substrates distorts the enzyme and that *the binding energy is now stored in the enzyme. It is as if the enzyme contained an internal spring which would be compressed when the substrate binds.* This would keep the binding weak but the energy in the spring might then be used to increase the velocity.

### 5. Proofreading

Although enzymes may be very specific they do make mistakes. This is of particular concern for processes such as protein synthesis in which the correct amino acid is placed at each position in the sequence with an error rate that has been estimated for *E. coli* as only 1 in $10^4$. It would be impossible for an enzyme designed to attach valine to its specific transfer RNA to avoid attaching the smaller alanine if discrimination between the two were based solely on the Gibbs energy differences of binding.[103] However, it would be easier for an enzyme to exclude *larger* amino acids. This problem may be resolved by use of multistep screening.[104,105] For example, isoleucyl-tRNA synthetase (Chapter 29) does occasionally attach the smaller valine to the specific tRNA,$^{Ile}$ but when it does the enzyme in a "proofreading and editing" step hydrolyzes off the incorrect amino acid. The active site for this hydrolyzing activity, whether at a different place on the enzyme surface or created by a conformational change, may be able to exclude sterically the larger isoleucyl residue while acting on the valyl-tRNA. This editing mechanism for isoleucyl-tRNA synthetase was demonstrated directly in 1998 by X-ray crystallography on complexes of the enzyme with L-isoleucine and L-valine. Both substrates fit into the ATP-requiring synthetic site but neither isoleucine nor isoleucyl-tRNA will fit into the editing site which is located in an adjacent β-barrel domain.[104,105] Proofreading steps based on differing chemical properties as well as size can also be visualized.[103,106]

## D. Mechanisms of Catalysis

Kinetic studies tell how fast enzymes act but by themselves say nothing about *how* enzymes catalyze reactions. They do not give the **chemical mechanism** of catalysis, the step-by-step process by which a reaction takes place. Most of the individual steps involve the simultaneous breaking of a chemical bond and formation of a new bond. Consider a simple **displacement reaction**, that of a hydroxyl ion reacting with methyl iodide to give the products methanol and iodide ions.

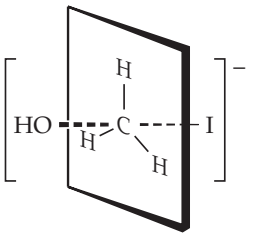



(9-76)

The reaction can be thought of as an "attack" by the $OH^-$ ion on the "back side" of the carbon atom of the methyl groups with a simultaneous displacement of the $I^-$.

### 1. The "Transition State"

A reaction such as that of Eq. 9-76 is not instantaneous, and at some time between that at which the reactants exist and that at which products have been formed the C–I bond will be stretched and partially broken and the new C–O bond will be partially formed. The structure at this point is not that of an ordinary compound and is energetically unstable with respect to both the reactant and the products. The intermediate structure of the least stability is known as the **transition state**. Although no one has actually seen a transition state structure, we might represent that for Eq. 9-76 as follows:




Transition state structure

The negative charge is distributed between the attacking $HO^-$ group and the departing iodide. The bonds to the central carbon atom are no longer tetrahedrally arranged but the C–H bonds lie in a single plane and the partial bonds to the OH and iodine atom lie at right angles to that plane.

It is useful in discussing a reaction mechanism to construct a **transition state diagram** in which Gibbs energy *G* is plotted against **reaction coordinate** (Fig. 9-15A). Energy *E* or enthalpy *H* may be plotted in the same way and authors frequently do not state whether *G*, *E*, or *H* is meant. The reaction coordinate is usually not assigned an exact physical meaning but represents the progress from reactants toward products. It is directly related to the extent to which an existing bond has been stretched and broken or a new one formed. The high energy point is the transition state. A somewhat more detailed idea of a transition state is obtained from a contour diagram such as that of Fig. 9-15B. Here, energy is plotted as a function of two distances, e.g., the lengthening C–I bond distance and the shortening C–O distance for Eq. 9-76. The path of minimum energy across the "saddle point" representing the transition state is indicated by the dashed line.

In a reaction coordinate diagram the difference in value of *G* between reactants and products is the

overall Gibbs energy change $\Delta G$ for the reaction, while the difference in $G$ between the transition state and reactants is $\Delta G^{\ddagger}$, the **Gibbs energy of activation**. The magnitude of $\Delta G^{\ddagger}$ represents the "energy barrier" to a reaction and largely determines the rate constant.

The diagrams in Fig. 9-15 are too simple because enzymatic reactions usually occur in several intermediate steps. There will be transition states for each step with valleys in between. The valleys correspond to intermediate species, which are sometimes very unstable. The passing from reactants to products in an enzymatic reaction can be likened to wandering through a series of mountain ranges of various heights and finally reaching the other side.

***Quantitative transition state theory.***[107–113] In the 1880s Arrhenius observed that the rate of chemical reactions varies with temperature according to Eq. 9-77 in a manner similar to the variation of an equilibrium constant with temperature (integrate Eq. 6-37 and compare). Here, $k$ is a first-order rate constant, the quantity $E_a$ is known as the **Arrhenius activation energy**, and the constant $A$ is referred to as the "pre-exponential factor" or the "frequency factor".

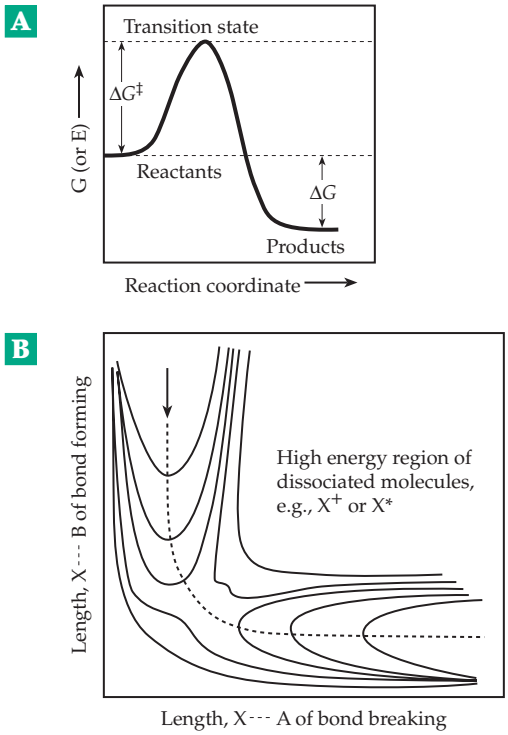



**Figure 9-15** (A) Transition state diagram illustrating Gibbs energy vs reaction coordinate for conversion of reactants to products in a chemical reaction. (B) Contour map of Gibbs energy vs interatomic bond distances for reaction B + X – A → B – X + A.

$$k = A e^{-E_a / RT} \tag{9-77}$$

The Arrhenius equation, together with studies of the effects of salts on reaction rates and observation of quantitative correlations between rates and equilibrium constants, suggested that a rate constant for a reaction might be a product of a constant term which is nearly independent of temperature and a constant $K^{\ddagger}$ which has the properties of an equilibrium constant for formation of the transition state. Eyring made this quantitative in 1935 with his "absolute rate theory"[112] according to which all transition states break down with a rate constant $\kappa k_B T/h$. Eyring reached this conclusion by assuming that *the rate of a chemical reaction is determined by the frequency of stretching of the bond that is being broken in the transition state*. To be more precise, it is the "normal-mode" oscillation of the transition state complex along the reaction coordinate.[109] This frequency $\nu$ was deduced by describing the vibrational energy as $h\nu$ (from quantum mechanics) and as $k_B/T$ (from classical mechanics) and setting them equal.

$$h\nu = k_B T \quad \text{and} \quad \nu = k_B T/h$$

$k_B$ is the Boltzmann constant
$h$ is Planck's constant. (9-78)

The right side of Eq. 9-78 is usually multiplied by a **transmission coefficient** $\kappa$, which may vary from 1 to 0.1 or even much less. However, for lack of any better value, $\kappa$ is usually assumed to be 1. From Eq. 9-78, at 25°C $\nu = 6.2 \times 10^{12}$ s$^{-1}$. This is the maximum rate for a chemical reaction of molecules in the transition state. This is the rate for a single molecule and must be multiplied by the concentration of the reacting substance X in the transition state. This concentration $[X]^{\ddagger}$ is determined by the equilibrium constant $K^{\ddagger} = [X]^{\ddagger}/[X]$. The velocity of the reaction becomes

$$\frac{-\mathrm{d}[X]}{\mathrm{dt}} = \nu = \frac{k_B T}{h} \cdot K^{\ddagger}[X] \tag{9-79}$$

The first-order rate constant is

$$k_1 = \frac{-\mathrm{d}[X]/\mathrm{dt}}{[X]} = K \frac{k_B T}{h} \cdot K^{\ddagger} \quad \mathrm{s}^{-1} \tag{9-80}$$

Since $\Delta G^{\ddagger} = -RT \ln K^{\ddagger}$, Eq. 9-80 can be rewritten as Eq. 9-81, in which $\Delta G^{\ddagger}$ is the Gibbs energy of activation.

$$k_1 \approx \frac{k_B T}{h} \cdot e^{-\Delta G^{\ddagger}/RT} \quad \mathrm{s}^{-1} \tag{9-81}$$

At 25°C, $\Delta G^{\ddagger}$ in kJ/mol, the following "practical" form of Eq. 9-81 can be written.

$$k_1 = 6.2 \times 10^{12}\, e^{-\Delta G^{\ddagger}/2.48}\ s^{-1}$$
$$\log k_1 = 12.79 - \Delta G^{\ddagger}/5.71$$
$$\text{and } \Delta G^{\ddagger} = 73.0 - 5.71 \log k_1$$

Equation 9-81 is approximate and a more correct statistical mechanical treatment is available.[113,114] See also comment on p. 288 about log $k$ (and log $k_1$) being unitless. Employing Eq. 6-14, we may expand Eq. 9-81 as follows:

$$k \approx \left(\frac{k_B T}{h} e^{\Delta S^{\ddagger}/R}\right) e^{-\Delta H^{\ddagger}/RT}\ s^{-1} \qquad (9\text{-}82)$$

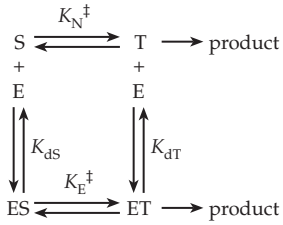
From this it appears that $\Delta H^{\ddagger} \quad E_a$, the Arrhenius activation energy. A more correct treatment gives $\Delta H^{\ddagger} = E_a - RT$ for reactions in solution. However, since $RT$ at 25°C is only 2.5 kJ mol$^{-1}$, the approximation that $\Delta H^{\ddagger} = E_a$ is often used. The preexponential term, in parentheses in Eq. 9-82, depends principally on $\Delta S^{\ddagger}$, the entropy change accompanying formation of the transition state. The quantities $\Delta G^{\ddagger}$, $\Delta H^{\ddagger}$, and $\Delta S^{\ddagger}$ are sometimes measured for enzymatic reactions but useful interpretations are difficult.

Equations 9-81 and 9-82 also show that a small decrease in transition state energy will give a large increase in rate; stabilization of the transition state by 5.7 kJ / mol (1.4 kcal / mol) will increase the rate 10-fold. If $\Delta G^{\ddagger}$ = 400 kJ/mol, as for a strong covalent single bond, $k_1 = 5 \times 10^{-58}$ s$^{-1}$ and the half-life $t_{1/2} = .693 / k_1 = 1.3 \times 10^{57}$ s ($4 \times 10^{49}$ years, greater than the ~$10^{10}$ years estimated age of the universe). If $\Delta G^{\ddagger}$ = 100 kJ / mol, $\log k_1 \approx -4.7$ s$^{-1}$, $k_1 = 1.9 \times 10^{-5}$, and $t_{1/2} \approx 10$ h. This is about the rate of a typical nonenzymatic "model" reaction at 25°C. If $\Delta G^{\ddagger}$ = 50 kJ / mol, $\log k_1 \approx 4$, as for a *fast* enzyme.

We can conclude that *enzymes make use of relatively small energy differences in catalyzing reactions*. The energies of numerous van der Waals interactions of hydrogen bonds and electrostatic attraction or repulsion of charges are sufficient. Nevertheless, we can see that anything that stabilizes the transition state (decreases $\Delta G^{\ddagger}$) will increase the rate of reaction. The role of a catalyst is to permit the formation of a transition state of lower energy (higher stability) than that for the uncatalyzed reaction. Stabilization of the transition state of a reaction by an enzyme suggests that the enzyme has a higher affinity for the transition state than it does for substrate or products, an idea that appears to have been expressed first by Haldane[41] and popularized by Pauling.[115]

> *I think that enzymes are molecules that are complementary in structure to the activated complexes of the reactions that they catalyze. The attraction of the enzyme molecule for the activated complex would thus lead to a decrease in its energy, and hence to a decrease in energy of activation of the reaction and to an increase in the rate of the reaction.*

***Transition state inhibitors.*** Suppose that a chemical reaction of a compound S takes place with rate constant $k_N$ through transition state T. Let the equilibrium constant for formation of T be $K_N^{\ddagger}$. Assume that an enzyme E can combine either with S with dissociation constant $K_{dS}$ or with the compound in its transition state structure T with dissociation constant $K_{dT}$ (Eq. 9-83).




(9-83)

If equilibrium is assumed for all four sets of double arrows it is easy to show that Eq. 9-84 holds.

$$K_E^{\ddagger} / K_N^{\ddagger} = K_{dS} / K_{dT} \qquad (9\text{-}84)$$

According to transition state theory, if the transmission coefficient $\kappa = 1$, T and ET will be transformed to products at the same rate. Thus, if the mechanisms of the nonenzymatic and enzymatic reactions are assumed the same, the ratio of maximum velocities for first-order transformation of ES and S will be given by Eq. 9-85. For some enzymes the ratio

$$k_E / k_N = K_E^{\ddagger} / K_N^{\ddagger} = K_{dS} / K_{dT} \qquad (9\text{-}85)$$

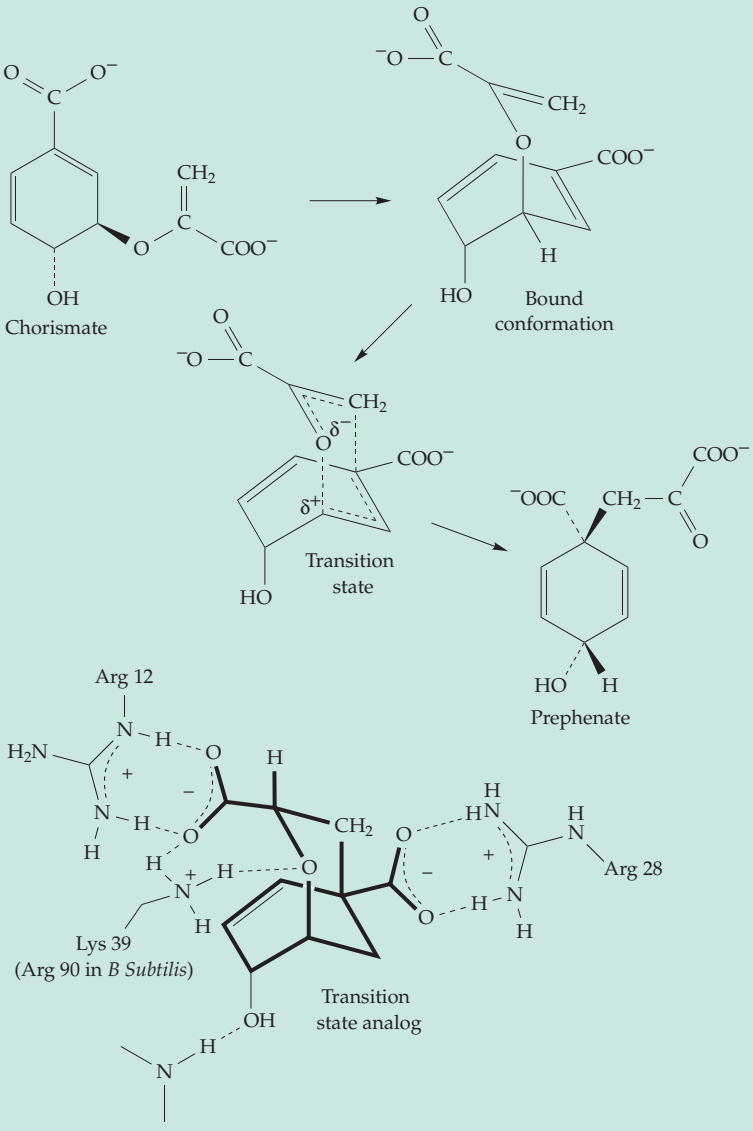
$k_E / k_N$ may be $10^8$ or more. Thus, if $K_{dS} \approx 10^{-3}$ the constant $K_{dT}$ would be $\approx 10^{-11}$. The enzyme would be expected to bind the transition state structure (T) $10^8$ times more tightly than it binds S.

The foregoing reasoning suggests that if structural analogs of T could be found for a particular reaction, they too might be very tightly bound—more so than ordinary substrate analogs.[116] Wolfenden[117] listed a series of compounds which may be **transition state inhibitors** of this type and many others have been described.[117a] Nevertheless, very tight, reversible binding is no *proof* that an inhibitor is actually binding to the transition state structure of the enzyme. For example, an inhibitor may bind to a conformational form of the enzyme that lies off to a side rather than directly on the reaction pathway for catalysis.

***Describing the transition state.*** How can we describe the structure of the transition state when we have no direct knowledge of its structure? We can try

## BOX 9-E CHORISMATE MUTASE

Bacteria, fungi, and plants convert products of glucose metabolism into the aromatic amino acids phenylalanine, tyrosine, and tryptophan by a complex sequence of reactions that is described in Chapter 25. One of these reactions, the conversion of **chorismate** into **prephenate,** is unique among enzyme-catalyzed reactions. It is a Claisen rearrangement that occurs readily *without catalysis* in an aqueous solution. However, the rate is increased over one million-fold by the enzyme **chorismate mutase**. The enzyme from *Bacillus subtilis* is a trimer of identical, small 127-residue subunits. Many studies, both of the nonenzymatic and enzymatic reactions, have suggested a "pericyclic" mechanism involving a polar chairlike transition state whose presumed structure is shown in the accompanying equation.[a–f] The transition state analog, whose structure is also depicted, is a powerful inhibitor. X-ray structural studies show that this inhibitor binds in the conformation shown with a number of interactions with polar side chain groups.[d,g] Some of these interactions are shown for the *E. coli* enzyme in the figure;[g] they are different in enzymes from *B. subtilis*[d] and yeast.[h] However, in no case are there groups in obvious positions to serve as acid–base catalysts and the maximum velocity $V_{max}$ is independent of pH.[b] NMR studies show that the bound product prephenate displays large shifts in some $^{13}C$ resonances, indicating strong interaction with the polar side chains of the protein.[c] Theoretical calculations also suggested that these interactions, especially those with Arg 90 and Glu 78 of the *B. subtilis* enzymes,[f–l] help to stabilize the partial charge separation in the polar transition state. Studies of mutant enzymes have confirmed that the charged Arg 90 (Lys 39 for *E. coli*) side chain is essential.[k]





[a] Gray, J. V., Eren, D., and Knowles, J. R. (1990) *Biochemistry* **29**, 8872–8878

[b] Turnbull, J., Cleland, W. W., and Morrison, J. F. (1991) *Biochemistry* **30**, 7777–7782

[c] Rajagopalan, J. S., Taylor, K. M., and Jaffe, E. K. (1993) *Biochemistry* **32**, 3965–3972

[d] Chook, Y. M., Gray, J. V., Ke, H., and Lipscomb, W. N. (1994) *J. Mol. Biol.* **240**, 476–500

[e] Gray, J. V., and Knowles, J. R. (1994) *Biochemistry* **33**, 9953–9959

[f] Lyne, P. D., Mulholland, A. J., and Richards, W. G. (1995) *J. Am. Chem. Soc.* **117**, 11345–11350

[g] Lee, A. Y., Karplus, P. A., Ganem, B., and Clardy, J. (1995) *J. Am. Chem. Soc.* **117**, 3627–3628

[h] Xue, Y., Lipscomb, W. N., Graf, R., Schnappauf, G., and Braus, G. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10814–10818

[i] Wiest, O., and Houk, K. N. (1995) *J. Am. Chem. Soc.* **117**, 11628–11639

[j] Kast, P., Hartgerink, J. D., Asif-Ullah, M., and Hilvert, D. (1996) *J. Am. Chem. Soc.* **118**, 3069–3070

[k] Lin, S. L., Xu, D., Li, A., Rosen, M., Wolfson, H. J., and Nussinov, R. (1997) *J. Mol. Biol.* **271**, 838–845

[l] Ma, J., Zheng, X., Schnappauf, G., Braus, G., Karplus, M., and Lipscomb, W. N. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14640–14645

to predict the atomic coordinates of all of the atoms of the reacting substrates and of the protein. The transition state structure of the substrate must be between that of the last ES complex and the first EP complex formed. It should be similar to that of a transition state inhibitor. We can hope to obtain suitable X-ray or NMR structures and deduce an appropriate enzyme structure by molecular modeling. Theoretical calculations using quantum chemical methods are of value in this effort.[118,119]

Lengths of the breaking and forming bonds in a transition state are often estimated using **kinetic isotope effects** (KIEs) on the velocity of the reaction.[118–122] This is somewhat indirect. Nevertheless, by measuring these effects for a large number of substrates and several isotopes a picture of the transition state structure that is good enough to use as a model for design of an enzyme inhibitor may be obtained.[117a,121–124] Other physical measurements such as Raman difference spectroscopy of isotopically labeled inhibitors can also be of value.[125]

***Getting to the transition state.*** Since some bonds must lengthen and some shorten and bond angles must also change, it may not be simple for an enzyme to catalyze necessary steps prior to the rate-limiting breakdown of the transition state. Hydrogen bonds within the protein may need to be broken and reformed with new bonding partners. Nonpolar side chains may need to shift to give a packing arrangement that will allow the movement of nuclei that must occur in forming the transition state structure. The enzyme must not only stabilize the transition state but also hold the substrate and escort it into the transition state configuration.

The attacking hydroxyl ion in Eq. 9-76 carries a negative charge which becomes distributed more or less equally between the hydroxyl and the iodine atom in the transition state. In this state the $CH_3$ group carries a partial positive charge as well. To provide good complementarity between an enzyme and this transition state, a change in the initial charge distribution within the enzyme will also be required. If a positively charged group initially binds the $OH^-$ ion it may have to lose part of its charge and a group next to the iodide atom may have to gain positive charge. In the nonenzymatic reaction of Eq. 9-76 the ionic atmosphere provided by positive counterions in solution can continuously readjust to keep the negative charge effectively balanced at every step along the reaction coordinate and through the transition state. Within enzymes this adjustment may occur via redistribution of electrical charges within the polarizable network of internal hydrogen bonds. The enzyme structure must allow this. Because of the complexity of an enzymatic transition state it may be hard to compare it with the transition state of the corresponding nonenzymatic reaction. Effects of pressure[125a] and of temperature[125b] also raise doubts about the simple picture of Eq. 9-83. It also follows that tight binding of substrates in the ground state does not necessarily interfere with transition state stabilization.[126–129a]

## 2. Microscopic Reversibility

The important statistical mechanical principle of microscopic reversibility asserts that *the mechanism of any chemical reaction considered in the reverse direction must be exactly the inverse of the mechanism of the forward reaction*. A consequence of this principle is that if the mechanism of a reaction is known, that of the reverse reaction is also known. Furthermore, it follows that the *forward and reverse reactions catalyzed by an enzyme must occur at the same active site* on the enzyme and the transition state must be the same in both directions. The principle of microscopic reversibility is often useful when the likelihood of a given mechanism is being considered. If a mechanism is proposed for a reversible reaction in one direction the principle of microscopic reversibility will give an unambiguous mechanism for the reverse reaction. Sometimes this reverse mechanism will be chemically untenable and, recognizing this, the enzymologist can search for a better one.

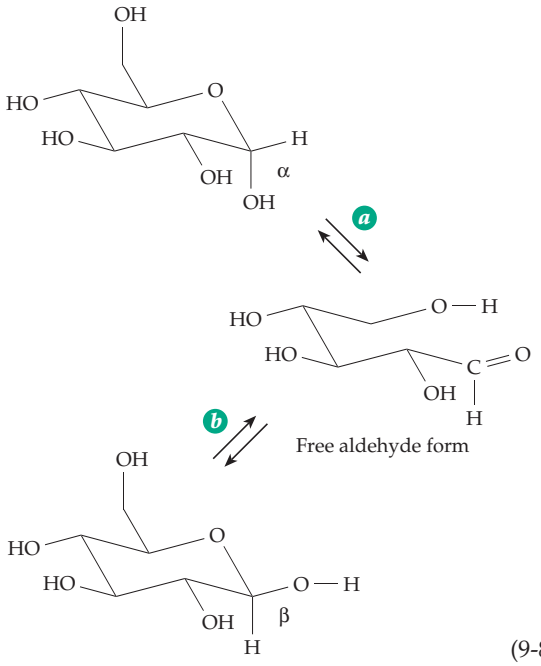
## 3. Acid and Base Catalysis

Many reactions that are promoted by enzymes can also be catalyzed by acids or bases or by both. An example is **mutarotation**, the reversible interconversion of the α- and β-anomeric forms of sugars (Eqs. 4-1 and 9-86). This reaction is catalyzed by a specific **mutarotase** and also by inorganic acids and bases. The frequently observed bell-shaped curve for the dependence of rate of catalysis on pH (e.g., Eqs. 9-55 to 9-57 and Fig. 9-8) also suggests participation of both protonated and unprotonated acid–base groups present in enzymes.

***Acidic and basic groups in enzymes.*** In this discussion the symbols HB or $H^+B$ will be used for acids and $B^-$ or B for their conjugate bases. Remember that *strong acids* have *low $pK_a$ values* and that the *conjugate bases formed from them are weak bases*. Likewise, *very weak acids have high $pK_a$ values and their conjugate bases are strong*. The following are among the $pK_a$ values that may be important in considering an enzyme mechanism. However, depending upon the environment of the side chain in a protein, these $pK_a$ values may fall substantially outside of the indicated ranges.

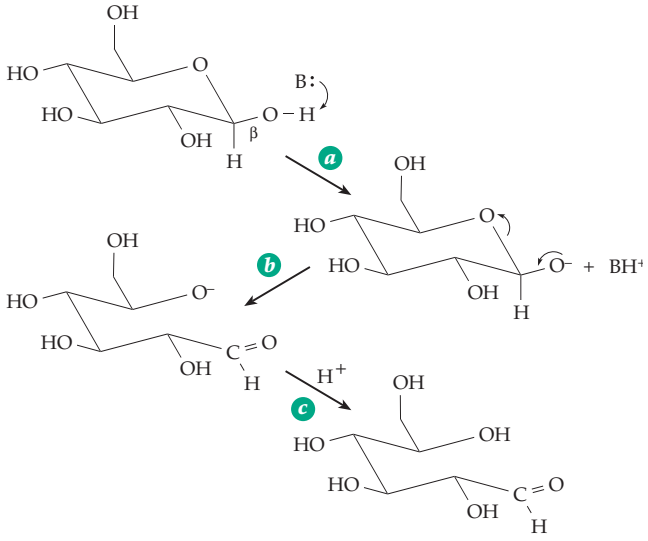
| | |
|---|---|
| −1.74 | $H_3O^+$ in 55.5 M water |
| ~3.6 | –COOH, terminal in peptide |
| 4.3–4.7 | –COOH in a glutamic or aspartic acid side chain |
| 6.4–7.0 | Imidazole (histidine) and phosphate ($–OPO_3H^–$) |
| 7.5–8.0 | $–NH_3^+$, terminal in peptide |
| 8.5–10 | –SH of cysteine side chains |
| 9.5–11 | Phenolic –OH of tyrosine |
| ~10.5 | $–NH_3^+$ of lysine side chain |
| ~13.6 | $–CH_2OH$ of serine side chain |
| 15.7 | free $HO^–$ in 55.5 M $H_2O$, 25°C |

***Acid–base catalysis of mutarotation.*** The mutarotation of glucose proceeds through the free aldehyde form as an intermediate (Eq. 9-86). A hydrogen atom is removed as $H^+$ from the oxygen at carbon 1 in step *a* and a proton (probably a different one) is transferred to the oxygen of the ring with cleavage of the O – C bond to the anomeric carbon atom. A similar process in reverse is required for step *b*. Transfers of hydrogen ions between atoms of oxygen, nitrogen, and sulfur atoms are a common feature of biochemical reactions. The bonds between hydrogen and O, N, and S atoms tend to be polarized strongly, leaving a partial positive charge on the hydrogen atoms. Consequently, the groups are weakly acidic and protons can be transferred from them relatively easily. It is reasonable to suppose that acid and base catalysis is related to these proton transfers.




(9-86)

***General base and general acid catalysis.*** Base-catalyzed mutarotation might be formulated as follows: A hydroxyl ion or some other base attacks the proton on the anomeric –OH group of the sugar, removing it to form an anion and the conjugate acid $BH^+$ (Eq. 9-87,




(9-87)

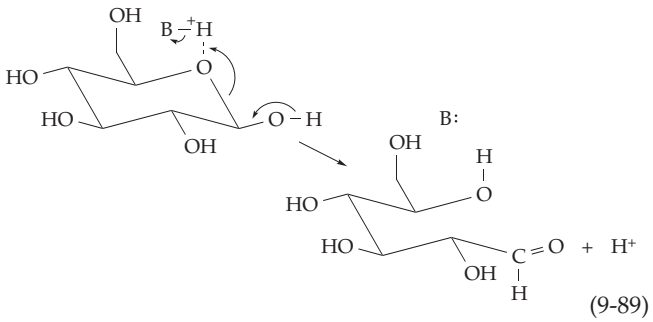
step *a*). The anion is isomerized to a second anion with the ring opened (step *b*). Addition of a proton (transfer of a proton from $H_3O^+$) produces the free aldehyde form of the sugar (step *c*).

The catalytic base B: might be $HO^–$ or a weaker base such as ammonia or even water. For reactions the rate is proportional only to the concentration of $OH^–$ and the presence of other weaker bases has no effect.[129b] Such catalysis is referred to as **specific hydroxyl ion catalysis**.[19] More commonly, the rate is found to depend both on $[OH^–]$ and on the concentration of other weaker bases. In such cases the apparent first-order rate constant ($k_{obs}$) for the process is represented by a sum of terms (Eq. 9-88). The term $k_{H_2O}$ is the rate in

$$k_{obs} = k_{H_2O} + K_{OH^–}[OH^–] + k_B[B] \qquad (9\text{-}88)$$

pure water and represents catalysis occurring by the action of water alone as either an acid or a base. The last two terms represent the contributions to the catalysis by $OH^–$ and by the other base, respectively. The term $k_B[B]$ represents **general base catalysis**.

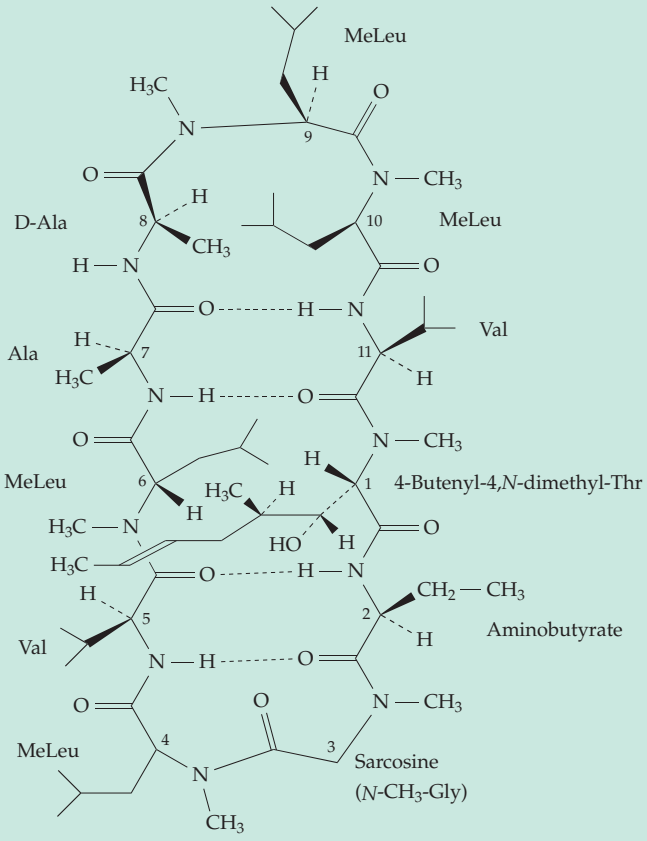
Catalysis of mutarotation by acids occurs if an acid donates a proton to the oxygen in the sugar ring as




(9-89)

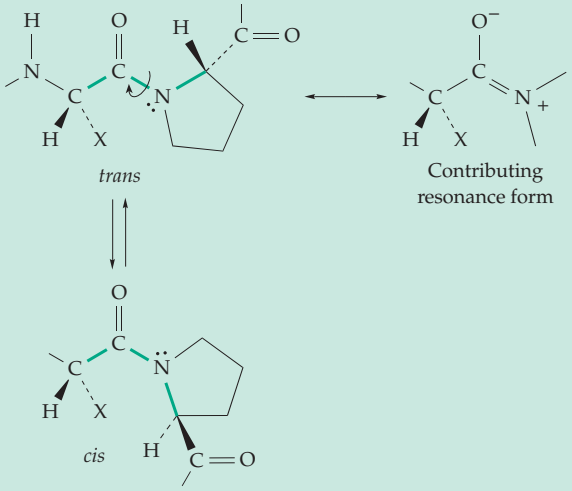
## BOX 9-F IMMUNOPHILINS AS ROTAMASES

Since its introduction into clinical use in about 1979 the immunosuppresant **cyclosporin** has been responsible for a revolution in human organ transplantation.[a–c] The exact mechanism of action in suppressing T-lymphocyte-mediated autoimmune responses is still not completely clear, but cyclosporin, a cyclic lipophilic peptide from a fungus, was found to bind to specific proteins that were named **cyclophilins**.[d] Human cyclophilin A is a 165-residue protein which associates, in the crystal form, as a decamer with five-fold rotational and dihedral symmetry.[e] This protein is also found in almost all other organisms and has a highly conserved sequence.[f] Another immunosuppresant, a synthetic molecule known only as **FK506**, has an action similar to that of cyclosporin but binds to distinctly different proteins, the FK506-binding proteins (FKBPs), of which the 107-residue FKBP-12 is the best known.[g–j] These proteins also bind another immunosuppresant, **rapamycin**, which, however, has different biological properties than does FK506. Both FK506 and rapamycin are macrocyclic compounds but with structures very different from that of cyclosporin.

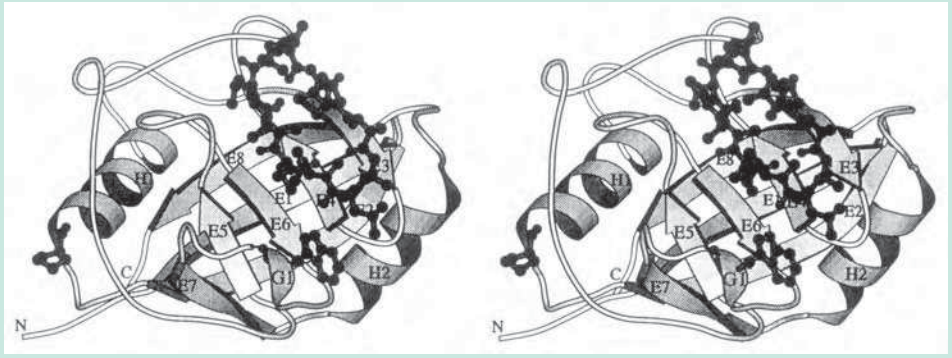



Cyclosporin A

It was a surprise to discover that all of the cyclophilins and FK506-binding proteins are **peptidyl prolyl *cis–trans* isomerases** or **rotamases**. They all catalyze the following simple and reversible reaction of a prolyl peptide linkage:




The transition state energy for the reaction is lowered by 33 kJ mol$^{-1}$ by cyclophilin A[k] and by 27 kJ mol$^{-1}$ by FKBP-12[g] with corresponding rate increases of ~ – 6 x 10$^5$- and 5 x 10$^4$-fold, respectively.



Cyclosporin (black) bound to human cyclophilin. From Pflügl *et al.*[e] Courtesy of J. N. Jansonius.

**BOX 9-F (continued)**

For both enzymes the maximum velocity is independent of pH over a wide region. For cyclophilin the rate is nearly independent of the nature of residue X in the foregoing structure but FKBP prefers a hydrophobic residue.

How do these enzymes work? One possibility would be for the protein to transfer a proton to the nitrogen atom of the proline ring. This would destroy the partial double-bond character of the amide linkage and allow free rotation about a single bond. The same thing could be accomplished if a nucleophilic group from the enzyme formed a covalent adduct with the C=O of the substrate. However, this seems unlikely because there is no suitably placed nucleophile. A third possibility is that the enzyme distorts the substrate and stabilizes the transition state using only noncovalent interactions.[g,h,l] This would account for the lack of pH dependence and observable solvent isotope effects. However, in human cyclophilin the guanidinium group of arginine 55 hydrogen bonds to the peptide C=O of the substrate's proline residue and could easily shift to place a guanidinium proton against the proline ring nitrogen effectively protonating it and permitting rotation about the peptide linkage.[m]

Do cyclosporin and FK506 act as transition state inhibitors? If so, we might learn something about the mechanism from the three-dimensional structures of the inhibited rotamases.[i,l,n,o] From these structures as well as those with simpler substrate analogs, it is seen that the substrate is desolvated and that many nonbonded (van der Waals) interactions stabilize the binding. Hydrogen bonding is also important.[p,q] The presence of distinct hydrogen bonds to the peptide NH and C=O groups on the N-terminal side of the substrate X-Pro linkage and less well-defined bonding on the terminal side suggests that the C-terminal part may be rotated while the enzyme holds the N-terminal part.[p] It can also be concluded that mechanisms are probably not identical for all rotamases.

Cyclophilins and FKBPs are large families of proteins. In *E. coli* there are two cyclophilin genes,[r,s] three encoding FK506-binding proteins, and one of a third family of bacterial rotamases (parvulins) that lack sequence similarity to the other two families.[t,u] Yeast contains at least five cyclophilins[v] and even more eukaryotic FKBPs are known.[j] There is abundant evidence that these proteins play an important role in protein folding *in vivo* (Chapter 29).


[a] Kahan, B. D. (1989) *N. Engl. J. Med.* **321**, 1725–1738
[b] Schreiber, S. L. (1991) *Science* **251**, 283–287
[c] High, K. P., Joiner, K. A., and Handschumacher, R. E. (1994) *J. Biol. Chem.* **269**, 9105–9112
[d] Fruman, D. A., Burakoff, S. J., and Bierer, B. E. (1994) *FASEB J.* **8**, 391–400
[e] Pflügl, G. M., Kallen, J., Jansonius, J. N., and Walkinshaw, M. D. (1994) *J. Mol. Biol.* **244**, 385–409
[f] Lippuner, V., Chou, I. T., Scott, S. V., Ettinger, W. F., Theg, S. M., and Gasser, C. S. (1994) *J. Biol. Chem.* **269**, 7863–7868
[g] Fischer, S., Michnick, S., and Karplus, M. (1993) *Biochemistry* **32**, 13830–13837
[h] Orozco, M., Tirado-Rives, J., and Jorgensen, W. L. (1993) *Biochemistry* **32**, 12864–12874
[i] Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) *J. Mol. Biol.* **229**, 105–124
[j] Lam, E., Martin, M. M., Timerman, A. P., Sabers, C., Fleischer, S., Lukas, T., Abraham, R. T., O'Keefe, S. J., O'Neill, E. A., and Wiederrecht, G. J. (1995) *J. Biol. Chem.* **270**, 26511–26522
[k] Eberhardt, E. S., Loh, S. N., Hinck, A. P., and Raines, R. T. (1992) *J. Am. Chem. Soc.* **114**, 5437–5439
[l] Kakalis, L. T., and Armitage, I. M. (1994) *Biochemistry* **33**, 1495–1501
[m] Zhao, Y., and Ke, H. (1996) *Biochemistry* **35**, 7362–7368
[n] Thériault, Y., Logan, T. M., Meadows, R., Yu, L., Olejniczak, E. T., Holzman, T. F., Simmer, R. L., and Fesik, S. W. (1993) *Nature (London)* **361**, 88–91
[o] Konno, M., Ito, M., Hayano, T., and Takahashi, N. (1996) *J. Mol. Biol.* **256**, 897–908
[p] Kern, D., Kern, G., Scherer, G., Fischer, G., and Drakenberg, T. (1995) *Biochemistry* **34**, 13594–13602
[q] Göthel, S. F., Herrler, M., and Marahiel, M. A. (1996) *Biochemistry* **35**, 3636–3640
[r] Clubb, R. T., Ferguson, S. B., Walsh, C. T., and Wagner, G. (1994) *Biochemistry* **33**, 2761–2772
[s] Edwards, K. J., Ollis, D. L., and Dixon, N. E. (1997) *J. Mol. Biol.* **271**, 258–265
[t] Rudd, K. E., Sofia, H. J., Koonin, E. V., Plunkett, G., III, Lazar, S., and Rouviere, P. E. (1995) *Trends Biochem. Sci.* **20**, 12–14
[u] Rahfeld, J.-U., Rücknagel, K. P., Stoller, G., Horne, S. M., Schierhorn, A., Young, K. D., and Fischer, G. (1996) *J. Biol. Chem.* **271**, 22130–22138
[v] Matouschek, A., Rospert, S., Schmid, K., Glick, B. S., and Schatz, G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6319–6323


shown in Eq. 9-89. Again, either **specific acid catalysis** (by $H_3O^+$) or **general acid catalysis** is possible.

Enzymes are not able to concentrate protons or hydroxyl ions to the point that *specific* base or acid catalysis would be effective. However, either general acid or general base catalysis can be accomplished by groups present in an enzyme in their normal states of protonation at the pH of the cell. Thus, if one of the enzymes must be dissociated to its conjugate base and a second must be protonated for reaction to take place (as in Eq. 9-54), it is reasonable to suppose that these two groups participate in acid–base catalysis. Of the acidic and basic groups present in proteins, the imidazole group of histidine would appear to be the

most ideal both for general base catalysis and, because both forms may exist in nearly equal amounts at pH 7, for general acid catalysis. However, carboxyl, thiol, lysyl, tyrosyl, and N-terminal amino groups are all thought to function in various enzymes.

The mutarotase of *E. coli* has a turnover number of $10^4$ $s^{-1}$. The plot of $-\log K_m$ vs pH indicates two p$K_a$ values in the free enzyme at 5.5 and 7.6, while the plot of log $V_{max}$ yields a single p$K_a$ of 4.75 for the ES complex.[130] The latter might represent a carboxyl group on an imidazole group in its conjugate base form. Why doesn't the group having p$K_a$ 7.6 in the free enzyme also show a p$K_a$ in the ES complex? Either the group has no catalytic function so that $EH_2S$ of Eq. 9-54 reacts to form products just as fast as does HES or the p$K_a$ is so strongly shifted by substrate binding that it is not detected in the log $V_{max}$ plot.

***The Brönsted relationships.*** The effectiveness of a specific base as a general base catalyst can usually be related to its basicity (p$K_a$) via the **Brönsted equation** (Eq. 9-90). Here, $k_B$ is defined by Eq. 9-88 and $G_B$ is a

$$\log k_B = \log G_B + \beta\,(\mathrm{p}K_a) \tag{9-90}$$

constant for a particular reaction. A similar equation (9-91) relates the constant $k_A$ for general acid catalysis

$$\log k_{HB} = \log G_A - \alpha(\mathrm{p}K_a) \tag{9-91}$$

to p$K_a$. These equations are linear Gibbs energy relationships similar to the ones discussed in Box 6-C. For the Brönsted equations to hold, the Gibbs energy of activation for the reaction must be directly related to the basicity or acidity of the catalyst.

The exponents β and α of Eqs. 9-90 and 9-91 measure the *sensitivity* of a reaction toward the basicity or acidity of the catalyst. It is easy to show that as β and α approach 1.0 general base or general acid catalysis is lost and that the rate becomes exactly that of specific hydroxyl ion or specific hydrogen ion catalysis.[131] As β or α approach zero basic or acidic catalysis is indetectable. Thus, general base or general acid catalysis is most significant when β or α is in the neighborhood of 0.5. Under these circumstances it is easy to see how a moderately weak basic group, such as the imidazole group of histidine, can be an unusually effective catalyst at pH 7.

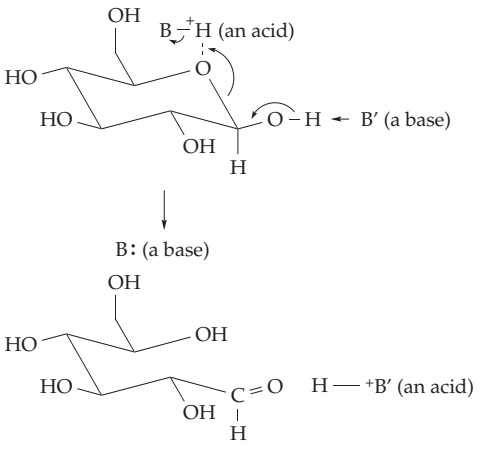
To determine α or β experimentally a plot of log $k_B$ or log $k_{HB}$ vs p$K_a$ (a **Brönsted plot**) is made and the slope is measured. Statistical corrections (Chapter 7) should be applied for dicarboxylic acids and for ammonium ions from which one of three protons may be lost from the nitrogen atom. General base or general acid catalysis implies an important feature of any mechanism for which it is observed, namely, that removal of a proton or addition of a proton is involved in the rate-determining step of a reaction.

***Concerted acid–base or tautomeric catalysis.*** A third possible type of catalysis requires that a base and an acid act *synchronously* to effect the breaking and formation of bonds in a single step. Thus, tetramethylglucose mutarotates very slowly in benzene containing either pyridine (a base) or phenol (an acid). However, when both pyridine and phenol are present, mutarotation is rapid. This suggested to Swain and Brown[132] a **concerted mechanism** (Eq. 9-92) in which both an acid and a base participate.

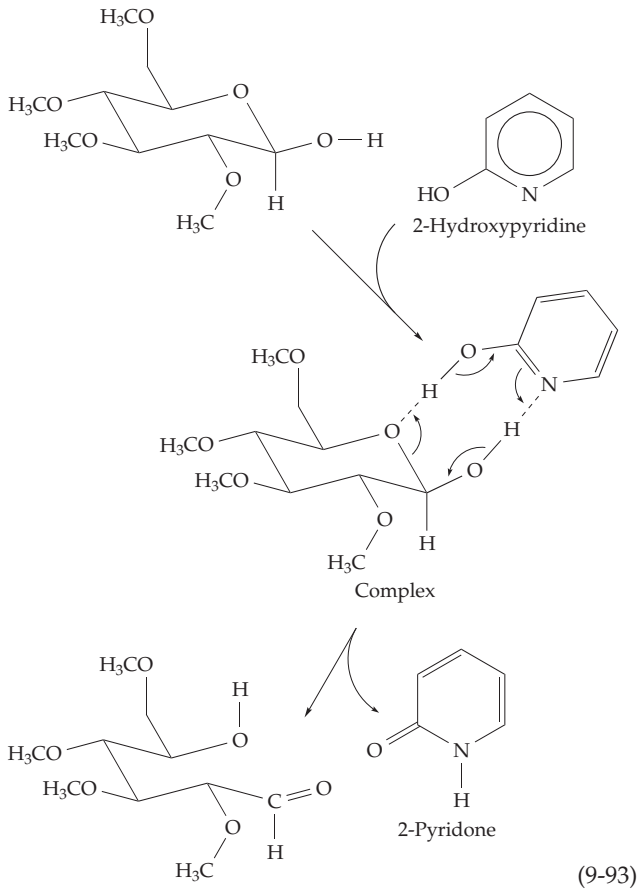
During the reaction shown, the acid $BH^+$ is converted to its conjugate base B and the base B′ to its conjugate acid $H^+B'$. It might seem that these agents, having been altered by the reaction, are not serving as true catalysts. However, a simple proton exchange will restore the original forms and complete the catalytic cycle. In aqueous solutions, water itself might act as the acid or the base or even both in concerted catalysis.

The original experimental evidence for concerted acid–base catalysis of the mutarotation in benzene is now considered unsound[133,134] and concerted acid–base catalysis has been difficult to prove for nonenzymatic reactions in aqueous solution. However, measurements of kinetic isotope effects seem to support Swain and Brown's interpretation.[135] Concerted acid–base catalysis by acetic acid and acetate ions may have been observed for the enolization of acetone[136] and it may be employed by enzymes.[136a]

Swain and Brown showed that a more effective catalyst for the mutarotation of sugars than a mixture of an acid and a base can be designed by incorporating the acidic and basic groups into *the same molecule*.[132,135] Thus, with 0.1 M tetramethylglucose in benzene solution, 0.001 M α-hydroxypyridine is 7000 times as effective a catalyst as a mixture of 0.001 M pyridine + 0.001 M phenol. Swain and Brown suggested the following completely concerted reaction mechanism for the **polyfunctional catalyst α-hydroxypyridine** in which the hydrogen-bonded complex formed (Eq. 9-93) is analogous to an enzyme–substrate complex. The product of the catalyst is 2-pyridone, a tautomer of




(9-92)




(9-93)

2-hydroxypyridine, with which it is in a rapid reversible equilibrium.

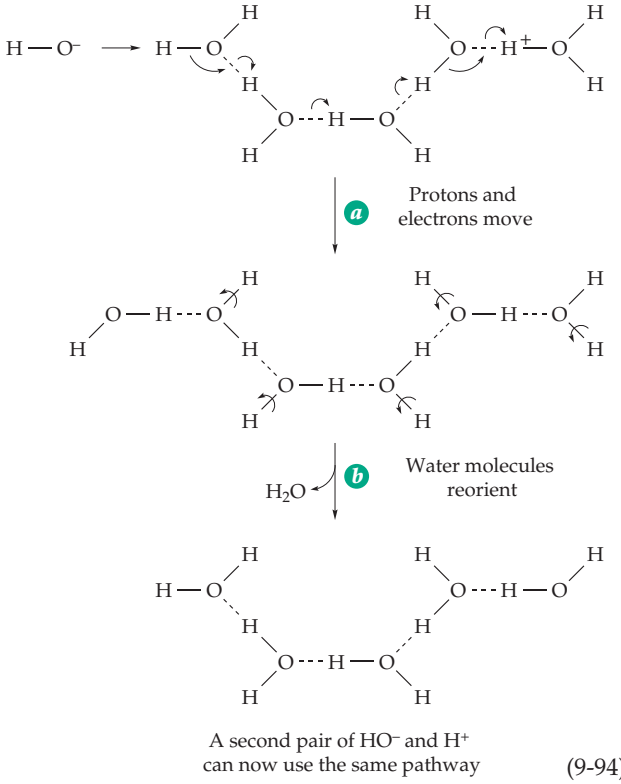
Rony called catalysis of the type illustrated in Eq. 9-93 tautomeric catalysis and suggested that its efficiency lies not simply in the close proximity of acidic and basic groups in the same molecule but also in the ability of the catalyst to repeatedly cycle between the two tautomeric states.[133] For an enzyme the tautomerization of the free catalyst could sometimes be rate determining (see Section A,4 on isomechanisms).

## 4. Hydrogen Bonding and the Transfer of Protons within Active Sites

An extensive network of hydrogen bonds runs throughout most proteins and may be especially complex within active sites. The network often runs through bound substrate molecules and immobilized water molecules in the active site cavity. This network arises in part because of the frequent occurrence of acidic and basic catalytic groups in active sites and by the fact that many substrates contain polar groups. The structure of an enzyme often seems to be more rigid in a complex with substrates or inhibitors than in the free state. Does this network of linked hydrogen bonds play a role in catalysis? If so, what? Wang suggested that rapid transfer of protons along rigidly and accurately held hydrogen bonds in the ES complex may be an essential feature of enzymatic catalysis.[137] This conclusion is supported by many more recent observations. A remaining question is whether this proton transfer can take place in such a way that the transition state barrier is lowered.

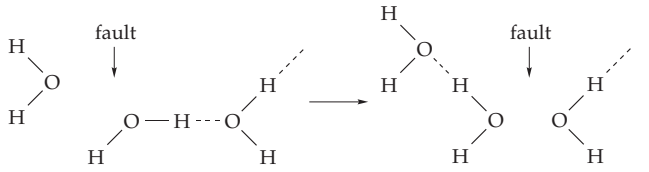
***Ultrafast proton transfer.*** The diffusion-controlled limit for second-order rate constants (Section A3) is ~ $10^{10}$ $M^{-1}$ $s^{-1}$. In 1956, Eigen, who had developed new methods for studying very fast reactions, discovered that protons and hydroxide ions react much more rapidly when present in a lattice of ice than when in solution.[138] He observed second-order rate constants of $10^{13}$ to $10^{14}$ $M^{-1}$ $s^{-1}$. These represent rates almost as great as those of molecular vibration. For example, the frequency of vibration of the OH bond in water is about $10^{14}$ $s^{-1}$. The latter can be deduced directly from the frequency of infrared light absorbed in exciting this vibration: Frequency ν equals wave number (3710 $cm^{-1}$ for –OH stretching) times *c*, the velocity of light ($3 \times 10^{10}$ cm $s^{-1}$).

This ultrafast transfer of protons can be explained as follows: The $OH^-$ ion and the proton, which is combined with a water molecule to form $H_3O^+$, are both hydrogen-bonded to adjacent water molecules. In ice a chain of hydrogen-bonded water molecules links the hydroxide and the hydrogen ions (Eq. 9-94). By synchronous movement of electron pairs from the $OH^-$ ion and from each of the water molecules in the




(9-94)

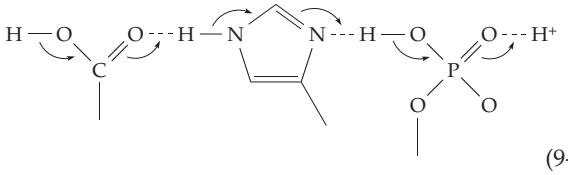
chain (as indicated by the little arrows) the neutralization can take place during the time of one molecular vibration (step *a*, Eq. 9-94). The positions of the oxygen atoms remain unchanged at the end of the reaction but the protons that were engaged in hydrogen bond formation have moved toward the left and are now attached to different oxygen atoms. The central chain of water molecules can be restored to its original state if each molecule rotates around one of its single bonds, swinging a hydrogen from right back to left (Eq. 9-94, step *b*). This must be a slower process which may occur one molecule at a time. The rotation of the leftmost water molecule would leave an empty space between two O-atoms, a "fault" in the H-bonded chain of the ice structure. This fault can be corrected by




(9-95)

rotation of the second molecule (Eq. 9-95); however, this would create a fault at the right of the second water molecule. This would induce rotation of the third molecule, etc., causing the fault to migrate from left to right across the entire chain of water molecules (Eq. 9-94, step *b*) and leaving the chain ready to function again in transferring another proton.

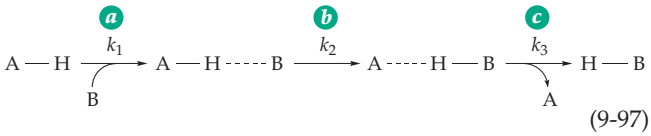
Because of the disorder present in ice, an array of water molecules such as that in Eq. 9-94 wouldn't revert to its exact original form in step *b*. However, active sites of enzymes are highly structured and proton transfers may occur with precision. For example, a synchronous shift of protons in an array of carboxylic acid, imidazolium, and phosphate groups can be envisioned readily (Eq. 9-96). The net effect of the process is to transfer a proton from one end of the chain to the other (as in Eq. 9-94) with facile tauto-



(9-96)

merization reactions providing the pathway. Such a pathway might be constructed by protein side chains to join the two sides of an active center promoting a concerted acid–base catalyzed reaction such as that of Eq. 9-92. Other tautomerization processes are possible within proteins if the existence of less stable minor tautomers of selected amide groups in the peptide backbone is allowed.[139] Nagle and Morowitz suggested a process similar to that of Eq. 9-94 but involving side chains of serine residues.[140]

***Proton transfer rates.*** Consider the reversible reaction of a proton acceptor B with acid H-A (Eq. 9-97). Eigen pointed out that the reaction will be fastest if the two reactants form a hydrogen-bonded complex (Eq. 9-97, step *a*).[138] The hydrogen bonding shortens the distance from the proton to B and allows for very rapid transfer of the proton from A to B within the




(9-97)

hydrogen-bonded complex (step *b*). The activation energy is close to zero. The complex dissociates in step *c* to form the products. The reactions are reversible, even though they have been indicated by unidirectional arrows.

The hydrogen-bonded complexes A–H—-B and A—-H–B can be formed readily if A and B are oxygen or nitrogen bases such as $-COO^-$, –OH, $-NH_2$, or imidazole. In such cases, as in ice, the interconversion of the two complexes (step *b* in Eq. 9-97) is very fast. The overall rate of the proton transfer will then be limited by diffusion of B and H–A together or by diffusion of A and H–B apart. It might seem that these two processes should also both be very fast. However, the rates will be determined by the concentration ratio [A–H—- B]/[A—- H–B]. If [A–H—- B] equals [A—- H – B] the rate of dissociation of A–H–B will be half what it is if nearly all of the complex is A—-H–B. If A–H—-B predominates by 1000 to 1, the rate will be slowed much more.

The equilibrium constant for the reaction of Eq. 9-97 will be:

$$K_{eq}^{AB} = K_a^{HA} / K_a^{HB} = k_f / k_r$$

or

$$\log K_{eq} = pK_a^{HA} - pK_a^{HB} = \log k_f - \log k_r \qquad (9\text{-}98)$$

Here $k_f$ and $k_r$ refer to the rates in the forward and reverse directions. If the $pK_a$'s of HA and HB are equal $k_f$ and $k_r$ will be the same, but if they are very far from equal the reaction will be slowed in one direction. If proton transfer is a step in an enzymatic reaction it may be slowed enough to become rate limiting.

The difference in $pK_a$ values between the proton donor and the proton acceptor in Eq. 9-97 can be expressed as the Gibbs energy change which at 25°C is equal to $\pm 5.71 \times \Delta pK_a$. This is often referred to as the **thermodynamic barrier** $\Delta G^\circ$ to a reaction and $\Delta G^\ddagger$ can be expressed as the sum of the thermodynamic barrier $\Delta G^\circ$ plus an **intrinsic barrier** $\Delta G^\ddagger_{int}$. For the proton transfer of Eq. 9-97 the intrinsic barrier (for step *b*) is thought to be near zero so that $\Delta G^\ddagger \approx 5.71\ \Delta pK_a$.

From this we can conclude that two p$K_a$ values can be as much as eight units apart and $\Delta G^{\ddagger}$ will still be less than 50 kJ / mol, low enough to permit rapid enzymatic reactions. However, for transfer of a proton from a C–H bond to a catalytic group, for example, to form an enolate ion for an aldol condensation (Chapter 13), the intrinsic barrier is known to be about 50 kJ / mol.[141] In this case, to allow rapid enzymatic reaction either the thermodynamic barrier must be very low, as a result of closely matching p$K_a$ values, or the enzyme must lower the intrinsic barrier. It may do both.

***Marcus theory.*** Discussion of intrinsic barriers is often approached using a quantitative theory proposed by Marcus.[142–145] It was first applied to electron transfer (Chapter 16) but has been used for a great variety of nonenzymatic and enzymatic reactions. As used by Gerlt and Gassman,[141] the Marcus formalism describes the reaction coordinate (Fig. 9-15A) as an inverted parabola whose shape is determined by the overall Gibbs energy change $\Delta G°$ and the intrinsic barrier $\Delta G^{\ddagger}_{int}$. The value of the Gibbs energy at any point on the curve is designated $G$ and the reaction coordinate $x$ is taken as 0 for the reactants and 1 for the products.

$$G = -4\,\Delta G^{\ddagger}_{int}\,(x - 0.5)^2 + \Delta G°\,(x - 0.5)$$
$$\text{Limits: } \Delta G°/4 \le \Delta G^{\ddagger}_{int} \le 4\,\Delta G° \qquad (9\text{-}99)$$

Differentiation of Eq. 9-99 yields the position of the transition state coordinate $x^{\ddagger}$ as follows:

$$x^{\ddagger} = 0.5 + \Delta G°/8\,\Delta G^{\ddagger}_{int} \qquad (9\text{-}100)$$

and it follows that

$$\Delta G^{\ddagger} = \Delta G^{\ddagger}_{int}\,(1 + \Delta G°/4\,\Delta G^{\ddagger}_{int})^2 \qquad (9\text{-}101)$$

***Diffusion-controlled dissociation of protons.*** The direct proton transfer between C-1 and C-2 during the action of sugar isomerases may seem puzzling. How can a highly mobile proton remain attached to a group in the enzyme for a millisecond or more instead of being transferred out to a solvent molecule? This can mean that the enzyme promotes the transfer of a hydride ion or of a hydrogen atom rather than a proton (see Chapter 13). If so, the observed proton exchange with solvent would be an unimportant side reaction. On the other hand, could the group in the enzyme that removes the proton be out of contact with the aqueous medium and thus able to hold onto the proton more tightly? In recent years, it has been recognized that neither of these explanations may be necessary. An imidazole group is a likely proton-carrying group at many active sites and it is thought that a proton cannot be expected to transfer out from an imidazolium group with a rate constant greater than ~$10^3$ $s^{-1}$.

The argument is as follows.[146] The rate of donation of a proton from $H_3O^+$ to imidazole (reverse of Eq. 9-102) is known to be diffusion controlled with a rate constant of 1.5 x $10^{10}$ $M^{-1}$ $s^{-1}$.

$$ImH^+ + H_2O \rightleftharpoons [ImH^+\cdots OH_2] \rightleftharpoons Im + OH_3^+$$
$$\uparrow$$
$$\text{Very fast diffusion controlled}$$
$$k \sim 1.5 \times 10^{10}\ M^{-1}\ s^{-1} \qquad (9\text{-}102)$$

The equilibrium constant for Eq. 9-102, calculated from the p$K_a$ of 7.0 for imidazole, is $10^{-7}$ M. Since $K_{eq}$ is also the *ratio of the overall rate constants for the forward and reverse reactions*, we see that for the forward reaction $k_f = 10^{-7}$ x 1.5 x $10^{10}$ = 1.5 x $10^3$ $s^{-1}$. This slow rate results from the fact that in the intermediate complex (in brackets in Eq. 9-102) the proton is on the imidazole group most of the time. For a small fraction of the time it is on the coordinated molecule $H_2O$ but reverts to being on the imidazole many times before the imidazole and $OH_3^+$ separate (see also Eqs. 9-97 and 9-98). Because of this unfavorable equilibrium within the complex, the diffusion-controlled rate of proton transfer from a protonated imidazole to water is far less than for proton transfer in the reverse direction.

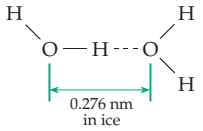
***Coupled proton transfers.*** Enzymatic reactions often require the transfer of two or more protons. They may move individually, one proton at a time, or as in Eq. 9-93 they may move synchronously in a **coupled** or concerted process. Such coupled movement is generally not possible for heavier nuclei.[147] However, studies of solvent isotope effects using a **proton inventory** technique[111,148,149] have provided evidence favoring coupled proton transfers for a variety of enzymes. Movement of protons along hydrogen-bonded paths, as well as electron transfer, may take place with some participation of **quantum mechanical tunneling**.[150–152] Coupling to vibrational modes of the hydrogen-bonded protons may provide **vibration-assisted tunneling**.[153–154d] These reactions are associated with unusually large kinetic isotope effects.

***Unusually strong hydrogen bonds.*** The strength of a hydrogen bond is thought to be directly related to the length, which is ordinarily taken to be the distance between the two surrounding heavier atoms (see Chapter 2, Section B,3). Hydrogen bonds are sometimes classified on the basis of N–H---O distances[155,156] as:

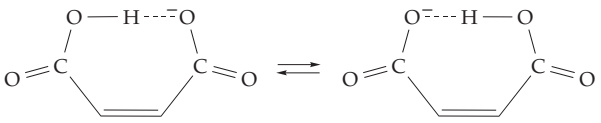
| | |
|---|---|
| <0.25 nm | very strong |
| 0.25–0.265 nm | strong |
| >0.28 nm | weak |
| >0.37 nm | no van der Waals contact but electrostatic interaction still occurs |

N–H---O distances may be a little longer than these.

The 0.276-nm hydrogen bonds in ice are regarded as moderately strong. However, if one of the oxygen atoms in an O–H---O hydrogen bond carries a negative charge, as in the maleate monoanion, it will be



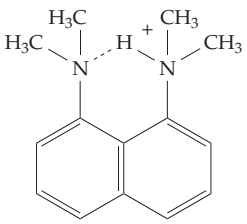

shorter.[156–158] Although the proton will be closer to one oxygen atom than to the other, it will be able to move between them by passing a very low transition state barrier. For this to occur the microscopic p$K_a$ values for the two groups (when they are protonated) must be similar; for the maleate anion they are identical.



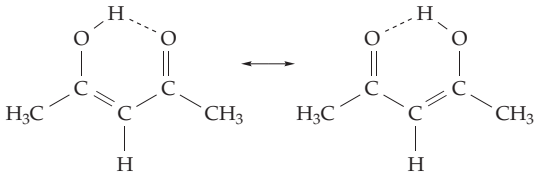
Monohydrogen maleate

The *strength* of a hydrogen bond can be measured for hydrogen-bonded complexes in the gas phase and range from 10–100 kJ/mol and even higher[155] for such complexes as $(FHF)^-$. It is more difficult to establish the strength in the liquid state or within the active site of an enzyme, but shorter hydrogen bonds are usually stronger than longer ones. Hydrogen bond distances in crystals of small molecules can be measured precisely by X-ray or neutron diffraction but greater uncertainty is present in distances within proteins. Lengths and probably strengths of hydrogen bonds in proteins can be measured by NMR methods.[158a,b] For example, solid-state NMR measurements on crystals of amino acids and other carboxylic acids have provided a plot of $^1H$ chemical shift vs hydrogen bond length.[159] For very short (0.24- to 0.25-nm) hydrogen bonds the $^1H$ chemical shift may be as great as 21 ppm. See also Fig. 3-30 and associated discussion. When dissolved in an $^1H_2O$–$^2H_2O$ mixture strongly hydrogen-bonded protons within a protein become enriched in $^1H$. The $^1H$ / $^2H$ ratio of the hydrogen-bonded protons provides another measure of the hydrogen bond strength.[158c]

A short hydrogen bond is also present in such cations as the following:



The p$K_a$ for dissociation of its proton is 12.3 and the hydrogen-bonded proton is probably located in the *center* of the bond with both amino groups sharing the charge.[160] Enols can also form unusually strong "resonance-assisted" hydrogen bonds:



The structure can be thought of as a resonance-stabilized enolate anion with a proton bound between the two oxygen atoms and equidistant from them.[156,157]

Do these short "**low-barrier hydrogen bonds**" have a special significance in enzymology? Proposals that they contribute to stabilization of transition states[161–166] have received some support[154,167] and aroused controversy.[158b,168–173] In later chapters we will examine specific enzymes in which low-barrier hydrogen bonds have been observed.

## 5. Covalent Catalysis

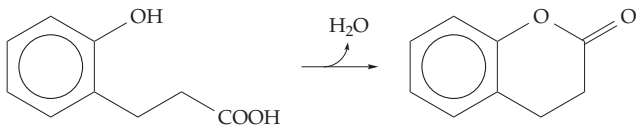
In addition to participating in acid–base catalysis, some amino acid side chains may enter into covalent bond formation with substrate molecules, a phenomenon that is often referred to as covalent catalysis.[174] When basic groups participate this may be called **nucleophilic catalysis**. Covalent catalysis occurs frequently with enzymes catalyzing nucleophilic displacement reactions and examples will be considered in Chapter 12. They include the formation of an acyl-enzyme intermediate by chymotrypsin (Fig. 12-11). Several of the coenzymes discussed in Chapters 14 and 15 also participate in covalent catalysis. These coenzymes combine with substrates to form reactive intermediate compounds whose structures allow them to be converted rapidly to the final products.

## 6. Proximity and Orientation

One of the earliest ideas about enzymes was that they simply brought reactants together and bound them side by side for a long enough time that the reactive groups might bump together and finally react. How important is this **proximity factor**? Page and Jencks estimated that rate enhancements by factors of $10^3$ or more may be expected solely from the loss in the entropy of two reactants when they are bound in close proximity on an enzyme surface.[107,175,176] In view of the large entropy decrease involved, *the enthalpy of*

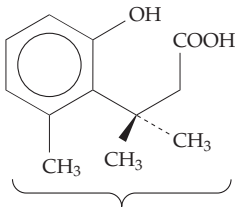
*binding must be high*, and if this explanation is correct the *binding of the substrates* to the enzyme provides much of the driving force for catalysis. Westheimer described this by stating that enzymes use the substrate-binding force as an **entropy trap**.[108] The losses of translational and rotational entropy, which Page and Jencks estimated as up to –160 to –210 J /deg /mol, overcome the unfavorable entropy of activation that is usual in bimolecular reactions.

How precise must the orientation of substrates be for rapid reaction?[107,177,177a] Compounds such as the acid shown in Eq. 9-103 form an internal ester (lactone) spontaneously with elimination of water.




(9-103)

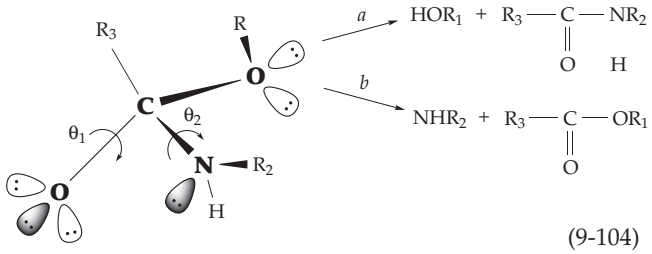
However, the following compound reacts at a rate over $10^{11}$ times that of the acid shown in Eq. 9-102. This is presumably because its conformation is highly restricted and the –COOH is constrained to frequently collide with the – OH group.[178,179] The three methyl




groups interdigitate and form a **trialkyl lock**. Orientation must also play a large role in enzymatic catalysis. As previously emphasized, enzymes often orient substrates precisely by formation of multiple hydrogen bonds. Small distortions by mutation or substitution of an essential metal by a different one can have very great effects. For example, because of an altered coordination pattern isocitrate dehydrogenases with $Ca^{2+}$ in the active site has a maximum velocity of catalysis only $2.5 \times 10^{-3}$ that with the normal $Mg^{2+}$-containing enzyme.[180] Bruice and associates concluded that enzymes must bring reacting groups into close proximity with orbitals of the reactants properly aligned in the ground state prior to moving to the transition state. They suggested that enzymes preorganize the enzyme-substrate complex into a **near attack conformation** in which the positions of reacting groups, the arrangements of hydrogen bonds, and the local dielectric constant in the active site, resemble those in the transition state. In this conformation the energy barrier to the transition state may be very low.[128-128b] Preorganization of the complex also acts to eliminate the slow components of solvent reorganization required for reaction in aqueous solutions.[128c,128d]

The necessity for reacting groups of substrates to collide with an orientation that allows productive interaction of electronic orbitals is often called a **stereoelectronic effect**. An example is the addition reaction of Eq. 9-74. The orbital of an unshared pair of electrons on the $HO^-$ ion must be perpendicular to the plane of the double bond. Furthermore, if the proton becomes attached to the adjacent carbon in a synchronous or concerted manner it must enter from the opposite side, as it does in Eq. 9-74.

In many biochemical reactions an alcohol or amine is eliminated from a tetrahedrally bonded intermediate as in Eq. 9-104. Deslongchamps proposed a stereoelectronic theory[181,182] according to which elimination of either NH—$R_2$ or –O—$R_1$ from this intermediate will depend upon the values of the torsion angles $\theta_1$ and $\theta_2$ (Eq. 9-104). The theory asserts that for rapid elimination of $OR_1$ (Eq. 9-104*a*), unshared electron pairs on *both* the $-O^-$ and N atoms must be antiperiplanar to the bond being broken (see Figure 2-2). If rotation around the C–O bond ($\theta_1$) occurs an orientation can be found in which an electron pair on each of the two oxygens will be antiperiplanar to the bond to $NHR_2$. From this orientation the latter will be eliminated (Eq. 9-104*b*). The theory has been supported by much experimental evidence involving reaction rates and product distribution among competing reactions of small conformationally restricted organic molecules.[181–184] Although more recent experiments[185] suggest that stereoelectronic factors may be of less significance than had been assumed, even a small decrease in transition state energy can be significant in an enzyme-catalyzed reaction. Enzymes may not only orient substrates in accord with stereoelectronic principles but also be able to promote conformational alterations in intermediates that allow them to take advantage of stereoelectronic factors. Examples are considered in Chapter 12.




(9-104)

## 7. The Microenvironment

A substrate bound at an active site may be in an environment quite different than that in an ordinary

aqueous solution. In fact, the protein often surrounds the substrate to the extent that *the protein is the solvent*. What kind of **microenvironment** does the protein provide and can it assist in the catalytic process? Charged and dipolar groups of the protein provide an electrostatic field that provides part of the binding energy and which may also assist in catalysis, a concept expressed by Quastel as early as 1926.[183a] The ends of protein helices often seem to point at active centers.[186,187] In many cases more than one N terminus with its positive electrical potential (see Chapter 2) or more than one C terminus (negative potential) of a helix point to an active site. These helix dipoles may be important in stabilizing transition state structures[188] or in altering p$K_a$ values of functional groups.[186] Fluctuations in charge distribution within a protein and in the hydrogen-bonding pattern of the protein with solvent and substrate may also be important.[189]

For some enzymatic reactions a transition state will be favored by a medium of very low dielectric constant and a correctly constructed active site can provide just such a surrounding.[189a] Hydrophobic groups may be packed around a site where an ion pair or other ionic interaction between enzyme and substrate occurs increasing the strength of that interaction. Conformational changes may enhance such effects. The dehydration of polar groups that must often occur upon binding of substrate may make these groups more reactive.[187,190,191] It has been suggested that the substantial volume changes ($\Delta V^{\ddagger}$) that sometimes occur during formation of transition states for enzyme-catalyzed reactions may result largely from changes in hydration of groups on the enzyme surface and that these changes may play an important role in catalysis.

In the past most enzymologists tacitly assumed that the external medium in which enzymes act must be aqueous. However, many enzymes function well in media containing largely hydrocarbons. Enzymes in a dry, powdered form have been suspended directly in organic solvents.[192,193] Under these conditions, enzymes may contain only tightly bound "structural" water, together with less than one equivalent of a monolayer of water outside the protein. The enzymes often remain active and are able to catalyze reactions with an altered substrate specificity as well as different reactions overall.

## 8. Strain and Distortion

The fact that enzymes appear to bind their substrates in such a way as to surround and immobilize them means that something other than the kinetic energy of the substrate is needed to provide energy for the ES complex to pass over the transition state barrier. What is the source of this activation energy? As with nonenzymatic reactions, it must come ultimately from the translational energy of solvent or solute molecules bombarding the complex. Can enzymes act as "energy funnels" that effectively channel kinetic energy from spots on the enzyme surface to the active site?[194] This could either be through strictly mechanical movement or through induction by fluctuations of the electrical field.[195]

It is often suggested that enzymes assist in catalysis by distorting bond lengths or angles away from their normal values. If the distorted structure were closer to the transition state geometry than the undistorted structure, catalysis would be assisted (see *lysozyme*, Chapter 12). However, Levitt[196] concluded (see also Fersht[2]) that forces provided this way by a protein are small and that the protein would become distorted rather than a substrate. On the other hand, if binding of a substrate distorts the protein, could the resultant "stored energy" be used in some way to assist in catalysis? If binding of a substrate bends a "spring" (Section C,4), can the tension in the spring then be used to distort the enzyme to be more exactly complementary to the transition state? This concept has become popular.[2,113,197]

The amino acid side chains of proteins are usually well packed. However, neither the side chains nor the main chain are rigid and immobile. Some regions of the protein will contain empty spaces – packing defects. Lumry called these **mobile defects** because, as a result of fluctuations in side chain packing, they can move within a protein from one site to another for a considerable distance.[198] Much evidence, including X-ray studies at low temperatures,[199] supports the existence of many **conformational substates** in proteins. Some substates may bind substrates better than others and some may allow conversion to the transition state more readily than do others. Rotation of histidine rings, peptide linkages, – OH groups, or amide side chains[199a] may be required and has, in fact, been observed for some enzymes. Perhaps in one of the substates of an enzyme–substrate complex an especially favorable vibrational mode[200] leads the complex to the transition state. In the transition state the packing of side chains may be especially tight. A mobile defect present in the active site may have moved elsewhere. The binding of substrates to the protein in the transition state will also be tighter because of the conformational alterations that have occurred. The binding energy of the substrates is now being utilized to lower the transition state barrier. The substrate is literally squeezed into the transition state configuration.

Is it possible that the protein domains forming an active site act like **ferroelectric crystals**, which change their dipole moment in response to a change in electric field? The highly polarizable hydrogen-bonded network, the amide linkages, imidazole rings, guanidinium group, etc. of active sites may permit a flip-flop of the dipole moment as in domains of ferroelectric crystals.

In such crystals, e.g., those of the hydrogen-bonded $KH_2PO_4$, a 180° change of dipole-moment direction results from very small movements of heavy atoms together with larger movements of the hydrogen-bonded protons.[200a,b,c] Could a similar flip-flop in a protein domain be coupled to the passage of a substrate over the transition state barrier? I have not seen any discussion of this possibility, but the structure of protein domains would seem to allow it. One could also imagine that with a small change in the hydrogen-binding arrangement of protein groups an active site could become preorganized to favor a flip-flop along a hydrogen-bonded chain in a different direction for a subsequent step in a reaction sequence.

## 9. Why Oligomeric Enzymes?

As we have seen (Chapter 7), a large fraction of all proteins exist as dimers, trimers, and higher oligomers. Oligomeric proteins raise the osmotic pressure much less than would the same number of monomeric subunits and this may be crucial to a cell. Another advantage of oligomers may be reflected in the fact that active sites of enzymes are often at interfaces between two or more subunits. This may enhance the ability of enzymes to undergo conformational rearrangements that are required during their action, just as hemoglobin changes its oxygen affinity in concert with a change in intersubunit contacts (Fig. 7-25).

A curious observation is that crystals of the dimeric pig heart malate dehydrogenase bind only one molecule of the substrate $NAD^+$ per dimer tightly; the second $NAD^+$ is bound weakly.[201,202] Similar *anticooperativity* has been reported for other crystalline dehydrogenases[203] and various other enzymes. An intriguing idea is that anticooperativity in binding might reflect a cooperative action between subunits during catalysis. Suppose that only conformation A of a dehydrogenase bound reduced substrate and $NAD^+$, while conformation B bound NADH and oxidized substrate. If reduced substrate and $NAD^+$ were present in excess and if oxidized substrate were efficiently removed from the scene by further oxidation, the following cooperative events could occur in the mixed AB dimer. The subunit of conformation A would bind substrates, react, and be converted to conformation B. At the same time, because of the strong AB interaction, the subunit that was originally in conformation B would be converted back to A and would be ready to initiate a new round of catalysis. Since conformation A has a low affinity for NADH, dissociation of the reduced coenzyme, which is often the slow step in dehydrogenase action, would be facilitated.[204–207] Such a **reciprocating** or **flip-flop** mechanism, suggested first by Harada and Wolfe,[204] is attractive because it provides a natural basis for the existence of the many known dimeric enzymes that do not exhibit evident allosteric properties. Attempts to verify this idea have largely failed. However, recent crystallographic studies of both glyceraldehyde phosphate dehydrogenase[206,207] and thymidylate synthase[208] are consistent with the proposal. There is still a possibility that coordinated reciprocal changes in distribution of electrical charges in the two subunits may also be important.

## 10. Summary

It appears that enzymes exert their catalytic powers by **first** bringing together substrates and binding them in proper orientations at the active site. **Second**, they often provide acidic and basic groups in the proper orientations to promote proton transfers within the substrate. **Third**, groups within the enzyme (especially nucleophilic groups) may enter into covalent interaction with the substrates to form structures that are more reactive than those originally present in the substrate. **Fourth**, the protein often closes around the substrate to immobilize it and to hold it in an environment, often lacking water, which could impede catalysis. The enzyme is also probably able to make small readjustments of its structure to provide good complementarity to the substrate at every stage and especially to the transition state structure. **Fifth**, the enzyme may be able to induce strain or distortion in the substrate perhaps accompanying a conformational change within the protein. The following question is often asked: "Why are enzymes such large molecules?"[209] At least part of the answer is that an enzyme usually interacts, sometimes via special domains, with other proteins.[206] Another part of the answer is evident when we consider that the formation of a surface complementary to that of the substrate and possessing reasonable rigidity requires a complex geometry in the peptide backbone. In addition, the enzyme must provide functional groups at the proper places to enter into catalysis. It may require a certain bulk to provide a low dielectric medium. Finally, if essential conformational changes occur during the course of the catalysis, we can only be surprised that nature has succeeded in packing so much machinery into such a small volume.

# E. Classification of Enzymes

An official commission of the International Union of Biochemistry (IUB) has classified enzymes in the following six categories:[210]

1. *Oxidoreductases.* Enzymes catalyzing dehydrogenation or other oxidation and reduction reactions.