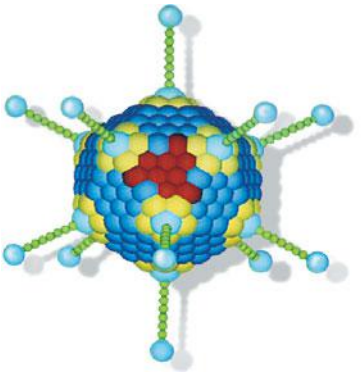


Introduction to Biochemistry

Anything found to be true of E. coli must also be true of elephants.

—Jacques Monod

Biochemistry is the discipline that uses the principles and language of chemistry to explain biology. Over the past 100 years biochemists have discovered that the same chemical compounds and the same central metabolic processes are found in organisms as distantly related as bacteria, plants, and humans. It is now known that the basic principles of biochemistry are common to all living organisms. Although scientists usually concentrate their research efforts on particular organisms, their results can be applied to many other species.

This book is called *Principles of Biochemistry* because we will focus on the most important and fundamental concepts of biochemistry—those that are common to most species. Where appropriate, we will point out features that distinguish particular groups of organisms.

Many students and researchers are primarily interested in the biochemistry of humans. The causes of disease and the importance of proper nutrition, for example, are fascinating topics in biochemistry. We share these interests and that's why we include many references to human biochemistry in this textbook. However, we will also try to interest you in the biochemistry of other species. As it turns out, it is often easier to understand basic principles of biochemistry by studying many different species in order to recognize common themes and patterns but a knowledge and appreciation of other species will do more than help you learn biochemistry. It will also help you recognize the fundamental nature of life at the molecular level and the ways in which species are related through evolution from a common ancestor. Perhaps future editions of this book will include chapters on the biochemistry of life on other planets. Until then, we will have to be satisfied with learning about the diverse life on our own planet.

We begin this introductory chapter with a few highlights of the history of biochemistry, followed by short descriptions of the chemical groups and molecules you will encounter throughout this book. The second half of the chapter is an overview of cell structure in preparation for your study of biochemistry.

Top: Adenovirus. Viruses consist of a nucleic acid molecule surrounded by a protein coat.

▲ **Friedrich Wöhler (1800–1882).** Wöhler was one of the founders of biochemistry. By synthesizing urea, Wöhler showed that compounds found in living organisms could be made in the laboratory from inorganic substances.

▲ **Some of the apparatus used by Louis Pasteur in his Paris laboratory.**

▲ **Eduard Buchner (1860–1917).** Buchner was awarded the Nobel Prize in Chemistry in 1907 "for his biochemical researches and his discovery of cell-free fermentation."

1.1 Biochemistry Is a Modern Science

Biochemistry has emerged as an independent science only within the past 100 years but the groundwork for the emergence of biochemistry as a modern science was prepared in earlier centuries. The period before 1900 saw rapid advances in the understanding of basic chemical principles such as reaction kinetics and the atomic composition of molecules. Many chemicals produced in living organisms had been identified by the end of the 19th century. Since then, biochemistry has become an organized discipline and biochemists have elucidated many of the chemical processes of life. The growth of biochemistry and its influence on other disciplines will continue in the 21st century.

In 1828, Friedrich Wöhler synthesized the organic compound urea by heating the inorganic compound ammonium cyanate.

$$NH_4(OCN) \xrightarrow{\text{Heat}} H_2N-\overset{\displaystyle O}{\overset{\|}{C}}-NH_2 \qquad (1.1)$$

This experiment showed for the first time that compounds found exclusively in living organisms could be synthesized from common inorganic substances. Today we understand that the synthesis and degradation of biological substances obey the same chemical and physical laws as those that predominate outside of biology. No special or "vitalistic" processes are required to explain life at the molecular level. Many scientists date the beginnings of biochemistry to Wöhler's synthesis of urea, although it would be another 75 years before the first biochemistry departments were established at universities.

Louis Pasteur (1822–1895) is best known as the founder of microbiology and an active promoter of germ theory. But Pasteur also made many contributions to biochemistry including the discovery of stereoisomers.

Two major breakthroughs in the history of biochemistry are especially notable—the discovery of the roles of enzymes as catalysts and the role of nucleic acids as information-carrying molecules. The very large size of proteins and nucleic acids made their initial characterization difficult using the techniques available in the early part of the 20th century. With the development of modern technology we now know a great deal about how the structures of proteins and nucleic acids are related to their biological functions.

The first breakthrough—identification of enzymes as the catalysts of biological reactions—resulted in part from the research of Eduard Buchner. In 1897 Buchner showed that extracts of yeast cells could catalyze the fermentation of the sugar glucose to alcohol and carbon dioxide. Previously, scientists believed that only living cells could catalyze such complex biological reactions.

The nature of biological catalysts was explored by Buchner's contemporary, Emil Fischer. Fischer studied the catalytic effect of yeast enzymes on the hydrolysis (breakdown by water) of sucrose (table sugar). He proposed that during catalysis an enzyme and its reactant, or substrate, combine to form an intermediate compound. He also proposed that only a molecule with a suitable structure can serve as a substrate for a given enzyme. Fischer described enzymes as rigid templates, or locks, and substrates as matching keys. Researchers soon realized that almost all the reactions of life are catalyzed by enzymes and a modified lock-and-key theory of enzyme action remains a central tenet of modern biochemistry.

Another key property of enzyme catalysis is that biological reactions occur much faster than they would without a catalyst. In addition to speeding up the rates of reactions, enzyme catalysts produce very high yields with few, if any, by-products. In contrast, many catalyzed reactions in organic chemistry are considered acceptable with yields of 50% to 60%. Biochemical reactions must be more efficient because by-products can be toxic to cells and their formation would waste precious energy. The mechanisms of catalysis are described in Chapter 5.

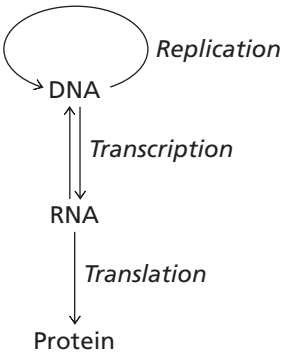
The last half of the 20th century saw tremendous advances in the area of structural biology, especially the structure of proteins. The first protein structures were solved in the 1950s and 1960s by scientists at Cambridge University (United Kingdom) led by

John C. Kendrew and Max Perutz. Since then, the three-dimensional structures of several thousand different proteins have been determined and our understanding of the complex biochemistry of proteins has increased enormously. These rapid advances were made possible by the availability of larger and faster computers and new software that could carry out the many calculations that used to be done by hand using simple calculators. Much of modern biochemistry relies on computers.

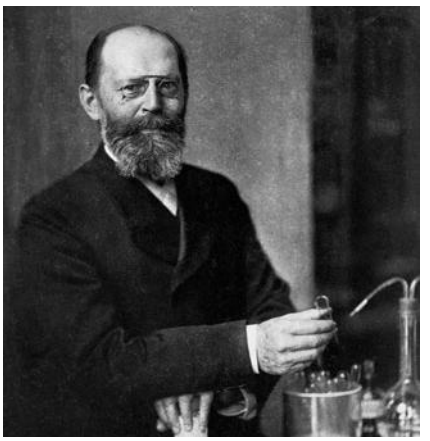
The second major breakthrough in the history of biochemistry—identification of nucleic acids as information molecules—came a half-century after Buchner's and Fischer's experiments. In 1944 Oswald Avery, Colin MacLeod, and Maclyn McCarty extracted deoxyribonucleic acid (DNA) from a pathogenic strain of the bacterium *Streptococcus pneumoniae* and mixed the DNA with a nonpathogenic strain of the same organism. The nonpathogenic strain was permanently transformed into a pathogenic strain. This experiment provided the first conclusive evidence that DNA is the genetic material. In 1953 James D. Watson and Francis H. C. Crick deduced the three-dimensional structure of DNA. The structure of DNA immediately suggested to Watson and Crick a method whereby DNA could reproduce itself, or replicate, and thus transmit biological information to succeeding generations. Subsequent research showed that information encoded in DNA can be transcribed to ribonucleic acid (RNA) and then translated into protein.

The study of genetics at the level of nucleic acid molecules is part of the discipline of molecular biology and molecular biology is part of the discipline of biochemistry. In order to understand how nucleic acids store and transmit genetic information, you must understand the structure of nucleic acids and their role in information flow. You will find that much of your study of biochemistry is devoted to considering how enzymes and nucleic acids are central to the chemistry of life.

As Crick predicted in 1958, the normal flow of information from nucleic acid to protein is not reversible. He referred to this unidirectional information flow from nucleic acid to protein as the *Central Dogma of Molecular Biology*. The term "Central Dogma" is often misunderstood. Strictly speaking, it does not refer to the overall flow of information shown in the figure. Instead, it refers to the fact that once information in nucleic acids is transferred to protein it cannot flow backwards from protein to nucleic acids.



▲ **Information flow in molecular biology.** The flow of information is normally from DNA to RNA. Some RNAs (messenger RNAs) are translated. Some RNA can be reverse transcribed back to DNA but according Crick's Central Dogma of Molecular Biology the transfer of information from nucleic acid (e.g., mRNA) to protein is irreversible.



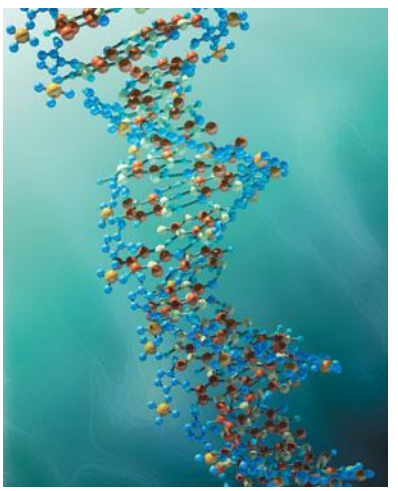
▲ **Emil Fischer (1852–1919).** Fischer made many contributions to our understanding of the structures and functions of biological molecules. He received the Nobel Prize in Chemistry in 1902 "in recognition of the extraordinary services he has rendered by his work on sugar and purine synthesis."

1.2 The Chemical Elements of Life

Six nonmetallic elements—carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur—account for more than 97% of the weight of most organisms. All these elements can form stable covalent bonds. The relative amounts of these six elements vary among organisms. Water is a major component of cells and accounts for the high percentage (by weight) of oxygen. Carbon is much more abundant in living organisms than in the rest of the universe. On the other hand, some elements, such as silicon, aluminum, and iron, are very common in the Earth's crust but are present only in trace amounts in cells. In addition to the standard six elements (CHNOPS), there are 23 other elements commonly found in living organisms (Figure 1.1). These include five ions that are essential in all species: calcium (Ca^{2+}), potassium (K^{+}), sodium (Na^{+}), magnesium (Mg^{2+}), and chloride (Cl^{-}) Note that the additional 23 elements account for only 3% of the weight of living organisms.

Most of the solid material of cells consists of carbon-containing compounds. The study of such compounds falls into the domain of organic chemistry. A course in organic chemistry is helpful in understanding biochemistry because there is considerable overlap between the two disciplines. Organic chemists are more interested in reactions that take place in the laboratory, whereas biochemists would like to understand how reactions occur in living cells.

Figure 1.2a shows the basic types of organic compounds commonly encountered in biochemistry. Make sure you are familiar with these terms because we will be using them repeatedly in the rest of this book.



▲ **DNA encodes most of the information required in living cells.**

IA												IIIA	IVA	VA	VIA	VIIA	0
1 H 1.008	IIA																2 He 4.003
3 Li 6.941	4 Be 9.012											5 B 10.81	6 C 12.01	7 N 14.01	8 O 16.00	9 F 19.00	10 Ne 20.18
11 Na 22.99	12 Mg 24.31	IIIB	IVB	VB	VIB	VIIB	VIIIB			IB	IIB	13 Al 26.98	14 Si 28.09	15 P 30.97	16 S 32.07	17 Cl 35.45	18 Ar 39.95
19 K 39.10	20 Ca 40.08	21 Sc 44.96	22 Ti 47.87	23 V 50.94	24 Cr 52.00	25 Mn 54.94	26 Fe 55.85	27 Co 58.93	28 Ni 58.69	29 Cu 63.55	30 Zn 65.39	31 Ga 69.72	32 Ge 72.61	33 As 74.92	34 Se 78.96	35 Br 79.90	36 Kr 83.80
37 Rb 85.47	38 Sr 87.62	39 Y 88.91	40 Zr 91.22	41 Nb 92.91	42 Mo 95.94	43 Tc (98)	44 Ru 101.1	45 Rh 102.9	46 Pd 106.4	47 Ag 107.9	48 Cd 112.4	49 In 114.8	50 Sn 118.7	51 Sb 121.8	52 Te 127.6	53 I 126.9	54 Xe 131.3
55 Cs 132.9	56 Ba 137.3	57 * La 138.9	72 Hf 178.5	73 Ta 180.9	74 W 183.8	75 Re 186.2	76 Os 190.2	77 Ir 192.2	78 Pt 195.1	79 Au 197.0	80 Hg 200.6	81 Tl 204.4	82 Pb 207.2	83 Bi 209.0	84 Po (209)	85 At (210)	86 Rn (222)
87 Fr (223)	88 Ra (226)	89** Ac (227)	104 Rf (261)	105 Db (262)	106 Sg (263)	107 Bh (264)	108 Hs (265)	109 Mt (268)	110 (269)	111 (272)	112 (277)	113	114 (285)	115	116 (289)	117	118 (293)

58* Ce 140.1	59 Pr 140.9	60 Nd 144.2	61 Pm (145)	62 Sm 150.4	63 Eu 152.0	64 Gd 157.3	65 Tb 158.9	66 Dy 162.5	67 Ho 164.9	68 Er 167.3	69 Tm 168.9	70 Yb 173.0	71 Lu 175.0
90** Th 232.0	91 Pa 231	92 U 238.0	93 Np (237)	94 Pu (244)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (251)	99 Es (252)	100 Fm (257)	101 Md (258)	102 No (259)	103 Lr (262)

▲ **Figure 1.1**
Periodic Table of the Elements. The important elements found in living cells are shown in color. The red elements (CHNOPS) are the six abundant elements. The five essential ions are purple. The trace elements are shown in dark blue (more common) and light blue (less common).

The synthesis of RNA (transcription) and protein (translation) are described in Chapters 21 and 22, respectively.

KEY CONCEPT

More than 97% of the weight of most organisms is made up of only six elements: carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur (CHNOPS).

KEY CONCEPT

Living things obey the standard laws of physics and chemistry. No "vitalistic" force is required to explain life at the molecular level.

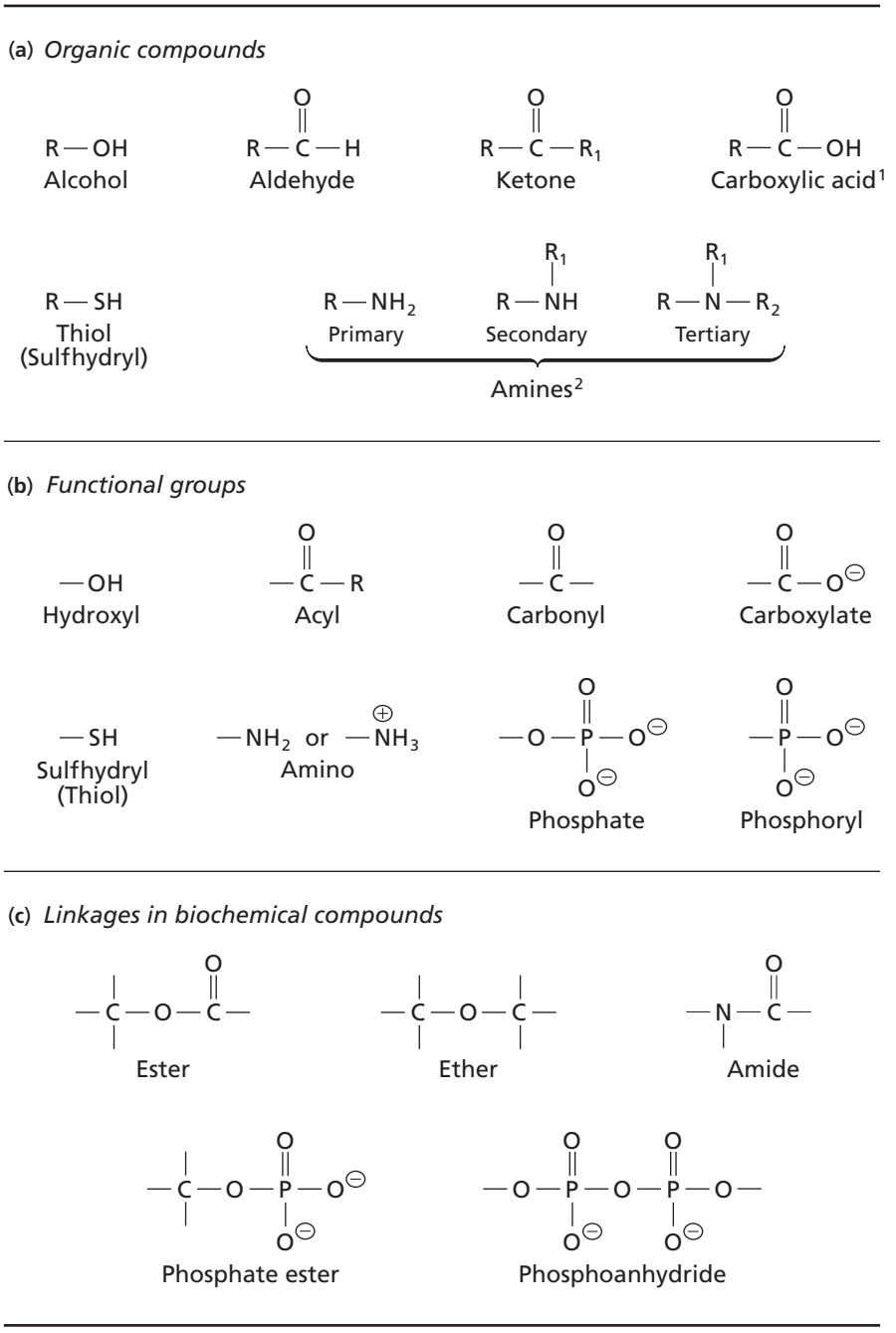
Biochemical reactions involve specific chemical bonds or parts of molecules called functional groups (Figure 1.2b). We will encounter several common linkages in biochemistry (Figure 1.2c). Note that all these linkages consist of several different atoms and individual bonds between atoms. We will learn more about these compounds, functional groups, and linkages throughout this book. Ester and ether linkages are common in fatty acids and lipids. Amide linkages are found in proteins. Phosphate ester and phosphoanhydride linkages occur in nucleotides.

An important theme of biochemistry is that the chemical reactions occurring inside cells are the same kinds of reactions that take place in a chemistry laboratory. The most important difference is that almost all reactions in living cells are catalyzed by enzymes and thus proceed at very high rates. One of the main goals of this textbook is to explain how enzymes speed up reactions without violating the fundamental reaction mechanisms of organic chemistry.

The catalytic efficiency of enzymes can be observed even when the enzymes and reactants are isolated in a test tube. Researchers often find it useful to distinguish between biochemical reactions that take place in an organism (*in vivo*) and those that occur under laboratory conditions (*in vitro*).

1.3 Many Important Macromolecules Are Polymers

In addition to numerous small molecules, much of biochemistry deals with very large molecules that we refer to as macromolecules. Biological macromolecules are usually a form of polymer created by joining many smaller organic molecules, or monomers, via condensation (removal of the elements of water). In some cases, such as certain carbohydrates, a single monomer is repeated many times; in other cases, such as proteins and nucleic acids, a variety of different monomers is connected in a particular order. Each monomer of a given polymer is added by repeating the same enzyme-catalyzed reaction.



◀ **Figure 1.2**
General formulas of **(a)** organic compounds, **(b)** functional groups, and **(c)** linkages common in biochemistry. R represents an alkyl group ($CH_3—(CH_2)_n—$).

[1] Under most biological conditions, carboxylic acids exist as carboxylate anions: $R—C(=O)—O^{\ominus}$

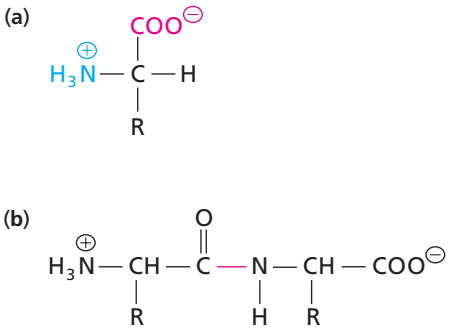
[2] Under most biological conditions, amines exist as ammonium ions: $R—NH_3^{\oplus}$, $R—NH_2^{\oplus}(R_1)$ and $R—NH^{\oplus}(R_1)—R_2$

Thus, all of the monomers, or residues, in a macromolecule are aligned in the same direction and the ends of the macromolecule are chemically distinct.

Macromolecules have properties that are very different from those of their constituent monomers. For example, starch is a polymer of the sugar glucose but it is not soluble in water and does not taste sweet. Observations such as this have led to the general principle of the hierarchical organization of life. Each new level of organization results in properties that cannot be predicted solely from those of the previous level. The levels of complexity, in increasing order, are: atoms, molecules, macromolecules, organelles, cells, tissues, organs, and whole organisms. (Note that many species lack one or more of these levels of complexity. Single-celled organisms, for example, do not have tissues and organs.) The following sections briefly describe the principal types of macromolecules and how their sequences of residues or three-dimensional shapes grant them unique properties.

In discussing molecules and macromolecules we will often refer to the **molecular weight** of a compound. A more precise term for molecular weight is **relative molecular mass** (abbreviated M_r). It is the mass of a molecule relative to one-twelfth (1/12) the mass of an atom of the carbon isotope ^{12}C. (The atomic weight of this isotope has been defined as exactly 12 atomic mass units. Note that the atomic weight of carbon shown in the Periodic Table represents the average of several different isotopes, including ^{13}C and ^{14}C.) Because M_r is a relative quantity, it is dimensionless and has no units associated with its value. The relative molecular mass of a typical protein, for example, is 38,000 (M_r = 38,000). The absolute molecular mass of a compound has the same magnitude as the molecular weight except that it is expressed in units called daltons (1 dalton = 1 atomic mass unit). The molecular mass is also called the molar mass because it represents the mass (measured in grams) of 1 mole, or 6.022×10^{23} molecules. The molecular mass of a typical protein is 38,000 daltons, which means that 1 mole weighs 38 kilograms. The main source of confusion is that the term "molecular weight" has become common jargon in biochemistry although it refers to relative molecular mass and not to weight. It is a common error to give a molecular weight in daltons when it should be dimensionless. In most cases, this isn't a very important mistake but you should know the correct terminology.

The relative molecular mass (M_r) of a molecule is a dimensionless quantity referring to the mass of a molecule relative to one-twelfth (1/12) the mass of an atom of the carbon isotope ^{12}C. Molecular weight (M.W.) is another term for relative molecular mass.



▲ **Figure 1.3**
Structure of an amino acid and a dipeptide. **(a)** Amino acids contain an amino group (blue) and a carboxylate group (red). Different amino acids contain different side chains (designated —R). **(b)** A dipeptide is produced when the amino group of one amino acid reacts with the carboxylate group of another to form a peptide bond (red).

KEY CONCEPT

Biochemical molecules are three-dimensional objects.

A. Proteins

Twenty common amino acids are incorporated into proteins in all cells. Each amino acid contains an amino group and a carboxylate group, as well as a side chain (R group) that is unique to each amino acid (Figure 1.3a). The amino group of one amino acid and the carboxylate group of another are condensed during protein synthesis to form an amide linkage, as shown in Figure 1.3b. The bond between the carbon atom of one amino acid residue and the nitrogen atom of the next residue is called a peptide bond. The end-to-end joining of many amino acids forms a linear polypeptide that may contain hundreds of amino acid residues. A functional protein can be a single polypeptide or it can consist of several distinct polypeptide chains that are tightly bound to form a more complex structure.

Many proteins function as enzymes. Others are structural components of cells and organisms. Linear polypeptides fold into a distinct three-dimensional shape. This shape is determined largely by the sequence of its amino acid residues. This sequence information is encoded in the gene for the protein. The function of a protein depends on its three-dimensional structure, or conformation.

The structures of many proteins have been determined and several principles governing the relationship between structure and function have become clear. For example, many enzymes contain a cleft, or groove, that binds the substrates of a reaction. This cavity contains the active site of the enzyme—the region where the chemical reaction takes place. Figure 1.4a shows the structure of the enzyme lysozyme that catalyzes the hydrolysis of specific carbohydrate polymers. Figure 1.4b shows the structure of the enzyme with the substrate bound in the cleft. We will discuss the relationship between protein structure and function in Chapters 4 and 6.

There are many ways of representing the three-dimensional structures of biopolymers such as proteins. The lysozyme molecule in Figure 1.4 is shown as a cartoon where the conformation of the polypeptide chain is represented as a combination of wires, helical ribbons, and broad arrows. Other kinds of representations in the following chapters include images that show the position of every atom. Computer programs that create these images are freely available on the Internet and the structural data for proteins can be retrieved from a number of database sites. With a little practice, any student can view these molecules on a computer monitor.

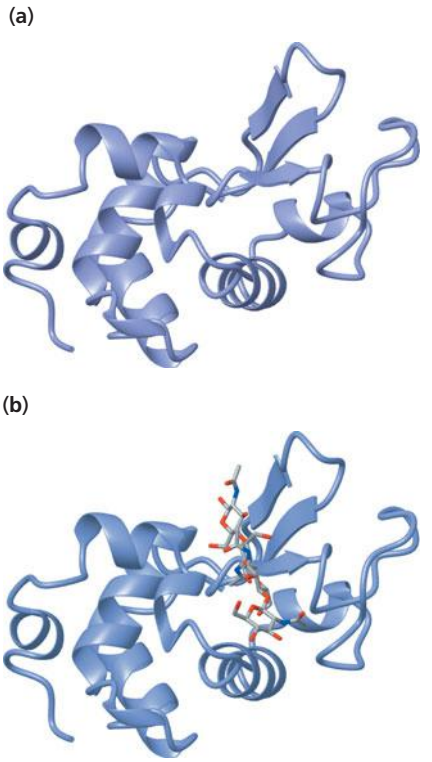
B. Polysaccharides

Carbohydrates, or saccharides, are composed primarily of carbon, oxygen, and hydrogen. This group of compounds includes simple sugars (monosaccharides) as well as their polymers (polysaccharides). All monosaccharides and all residues of polysaccharides contain several hydroxyl groups and are therefore polyalcohols. The most common monosaccharides contain either five or six carbon atoms.

Sugar structures can be represented in several ways. For example, ribose (the most common five-carbon sugar) can be shown as a linear molecule containing four hydroxyl groups and one aldehyde group (Figure 1.5a). This linear representation is called a Fischer projection (after Emil Fischer). In its usual biochemical form, however, the structure of ribose is a ring with a covalent bond between the carbon of the aldehyde group (C-1) and the oxygen of the C-4 hydroxyl group, as shown in Figure 1.5b. The ring form is most commonly shown as a Haworth projection (Figure 1.5c). This representation is a more accurate way of depicting the actual structure of ribose. The Haworth projection is rotated 90° with respect to the Fischer projection and portrays the carbohydrate ring as a plane with one edge projecting out of the page (represented by the thick lines). However, the ring is not actually planar. It can adopt numerous conformations in which certain ring atoms are out-of-plane. In Figure 1.5d, for example, the C-2 atom of ribose lies above the plane formed by the rest of the ring atoms.

Some conformations are more stable than others so the majority of ribose molecules can be represented by one or two of the many possible conformations. Nevertheless, it's important to note that most biochemical molecules exist as a collection of structures with different conformations. The change from one conformation to another does not require the breaking of any covalent bonds. In contrast, the two basic forms of carbohydrate structures, linear and ring forms, do require the breaking and forming of covalent bonds.

Glucose is the most abundant six-carbon sugar (Figure 1.6a on page 8). It is the monomeric unit of cellulose, a structural polysaccharide, and of glycogen and starch, which are storage polysaccharides. In these polysaccharides, each glucose residue is joined covalently to the next by a covalent bond between C-1 of one glucose molecule and one of the hydroxyl groups of another. This bond is called a glycosidic bond. In cellulose, C-1 of each glucose residue is joined to the C-4 hydroxyl group of the next residue (Figure 1.6b). The hydroxyl groups on adjacent chains of cellulose interact noncovalently creating strong, insoluble fibers. Cellulose is probably the most abundant biopolymer on Earth because it is a major component of flowering plant stems including tree trunks. We will discuss carbohydrates further in Chapter 8.



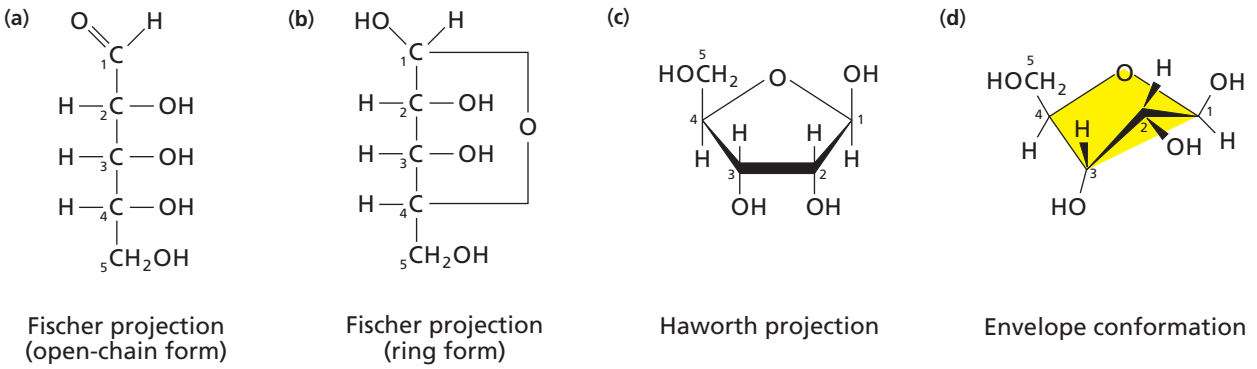
▲ **Figure 1.4 Chicken (*Gallus gallus*) eggwhite lysozyme. (a)** Free lysozyme. Note the characteristic cleft that includes the active site of the enzyme. **(b)** Lysozyme with bound substrate. [PDB 1LZC].

C. Nucleic Acids

Nucleic acids are large macromolecules composed of monomers called nucleotides. The term *polynucleotide* is a more accurate description of a single molecule of nucleic acid, just as *polypeptide* is a more accurate term than *protein* for single molecules composed of amino acid residues. The term *nucleic acid* refers to the fact that these polynucleotides were first detected as acidic molecules in the nucleus of eukaryotic cells. We

The rules for drawing a molecule as a Fischer projection are described in Section 8.1.

Conformations of monosaccharides are described in more detail in Section 8.3.



▲ **Figure 1.5**

Representations of the structure of ribose. (a) In the Fischer projection, ribose is drawn as a linear molecule. **(b)** In its usual biochemical form, the ribose molecule is in a ring, shown here as a Fischer projection. **(c)** In a Haworth projection, the ring is depicted as lying perpendicular to the page (as indicated by the thick lines, which represent the bonds closest to the viewer). **(d)** The ring of ribose is not actually planar but can adopt 20 possible conformations in which certain ring atoms are out-of-plane. In the conformation shown, C-2 lies above the plane formed by the rest of the ring atoms.

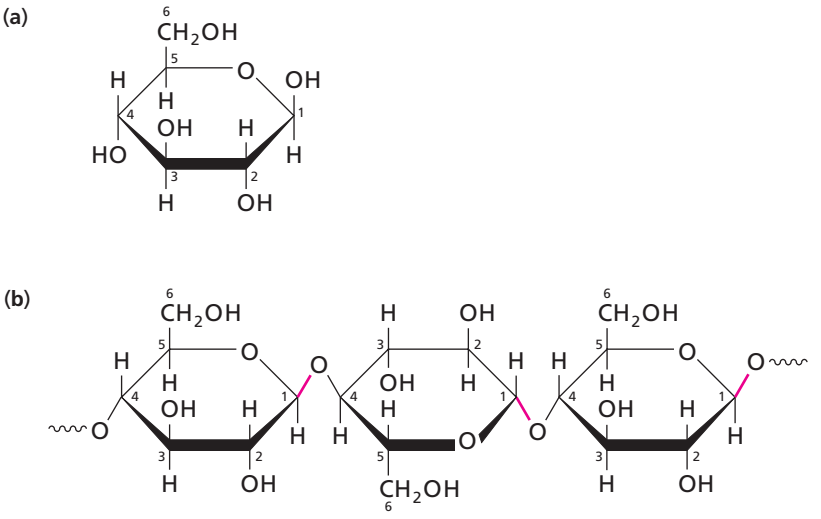
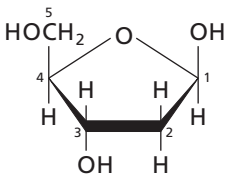


Figure 1.6 ►
Glucose and cellulose. (a) Haworth projection of glucose. **(b)** Cellulose, a linear polymer of glucose residues. Each residue is joined to the next by a glycosidic bond (red).

The structures of nucleic acids are described in Chapter 19.



▲ Figure 1.7
Deoxyribose, the sugar found in deoxyribonucleotides. Deoxyribose lacks a hydroxyl group at C-2.

now know that nucleic acids are not confined to the eukaryotic nucleus but are abundant in the cytoplasm and in prokaryotes that don't have a nucleus.

Nucleotides consist of a five-carbon sugar, a heterocyclic nitrogenous base, and at least one phosphate group. In ribonucleotides, the sugar is ribose; in deoxyribonucleotides, it is the derivative deoxyribose (Figure 1.7). The nitrogenous bases of nucleotides belong to two families known as purines and pyrimidines. The major purines are adenine (A) and guanine (G); the major pyrimidines are cytosine (C), thymine (T), and uracil (U). In a nucleotide, the base is joined to C-1 of the sugar, and the phosphate group is attached to one of the other sugar carbons (usually C-5).

The structure of the nucleotide adenosine triphosphate (ATP) is shown in Figure 1.8. ATP consists of an adenine moiety linked to ribose by a glycosidic bond. There are three phosphoryl groups (designated α, β, and γ) esterified to the C-5 hydroxyl group of the ribose. The linkage between ribose and the α-phosphoryl group is a phosphoester linkage because it includes a carbon and a phosphorus atom, whereas the β- and γ-phosphoryl groups in ATP are connected by phosphoanhydride linkages that don't involve carbon atoms (see Figure 1.2). All phosphoanhydrides possess considerable chemical potential energy and ATP is no exception. It is the central carrier of energy in living cells. The potential energy associated with the hydrolysis of ATP can be used directly in biochemical reactions or coupled to a reaction in a less obvious way.

The role of ATP in biochemical reactions is described in Section 10.7.

In polynucleotides, the phosphate group of one nucleotide is covalently linked to the C-3 oxygen atom of the sugar of another nucleotide creating a second phosphoester linkage. The entire linkage between the carbons of adjacent nucleotides is called a phosphodiester linkage because it contains two phosphoester linkages (Figure 1.9). Nucleic acids contain many nucleotide residues and are characterized by a backbone consisting of alternating sugars and phosphates. In DNA, the bases of two different polynucleotide strands interact to form a helical structure.

There are several ways of depicting nucleic acid structures depending on which features are being described. The ball-and-stick model shown in Figure 1.10 is ideal for showing the individual atoms and the ring structure of the sugars and the bases. In this case, the

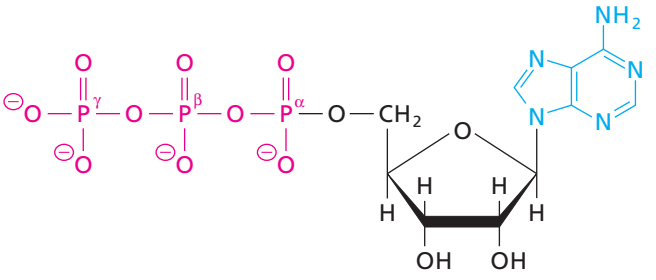
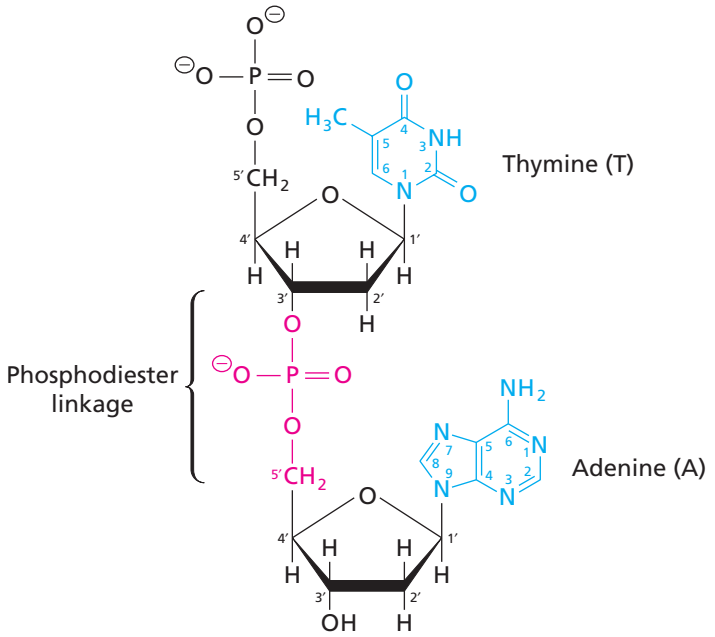


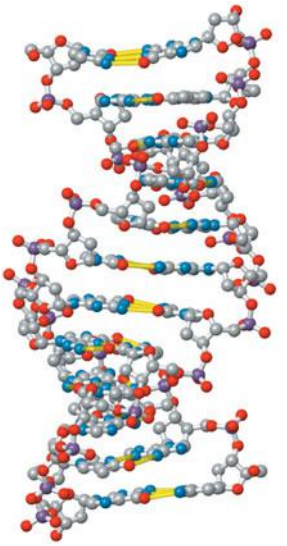
Figure 1.8 ►
Structure of adenosine triphosphate (ATP). The nitrogenous base adenine (blue) is attached to ribose (black). Three phosphoryl groups (red) are also bound to the ribose.



◀ **Figure 1.9**
Structure of a dinucleotide. One deoxyribonucleotide residue contains the pyrimidine thymine (top), and the other contains the purine adenine (bottom). The residues are joined by a phosphodiester linkage between the two deoxyribose moieties. (The carbon atoms of deoxyribose are numbered with primes to distinguish them from the atoms of the bases thymine and adenine.)

two helices can be traced by following the sugar–phosphate backbone emphasized by the presence of the purple phosphorus atoms surrounded by four red oxygen atoms. The individual base pairs are viewed edge-on in the interior of the molecule. We will see several other DNA models in Chapter 19.

RNA contains ribose rather than deoxyribose and it is usually a single-stranded polynucleotide. There are four different kinds of RNA molecules. Messenger RNA (mRNA) is involved directly in the transfer of information from DNA to protein. Transfer RNA (tRNA) is a smaller molecule required for protein synthesis. Ribosomal RNA (rRNA) is the major component of ribosomes. Cells also contain a heterogeneous class of small RNAs that carry out a variety of different functions. In Chapters 19 to 22, we will see how these RNA molecules differ and how their structures reflect their biological roles.



▲ **Figure 1.10**
Short segment of a DNA molecule. Two different polynucleotides associate to form a double helix. The sequence of base pairs on the inside of the helix carries genetic information.

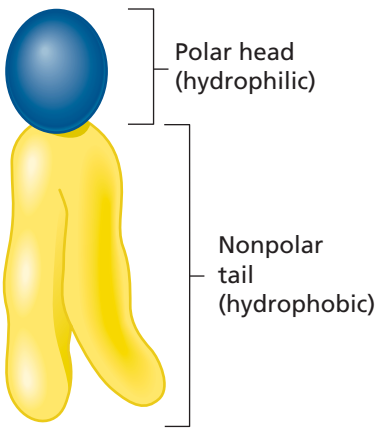
D. Lipids and Membranes

The term "lipid" refers to a diverse class of molecules that are rich in carbon and hydrogen but contain relatively few oxygen atoms. Most lipids are not soluble in water but they do dissolve in some organic solvents. Lipids often have a polar, hydrophilic (water-loving) head and a nonpolar, hydrophobic (water-fearing) tail (Figure 1.11). In an aqueous environment, the hydrophobic tails of such lipids associate while the hydrophobic heads are exposed to water, producing a sheet called a lipid bilayer. Lipid bilayers form the structural basis of all biological membranes. Membranes separate cells or compartments within cells from their environments by acting as barriers that are impermeable to most water-soluble compounds. Membranes are flexible because lipid bilayers are stabilized by noncovalent forces.

The simplest lipids are fatty acids—these are long-chain hydrocarbons with a carboxylate group at one end. Fatty acids are commonly found as part of larger molecules called glycerophospholipids consisting of glycerol 3-phosphate and two fatty acyl groups (Figure 1.12 on the next page). Glycerophospholipids are major components of biological membranes.

Other kinds of lipids include steroids and waxes. Steroids are molecules like cholesterol and many sex hormones. Waxes are common in plants and animals but perhaps the most familiar examples are beeswax and the wax that forms in your ears.

Membranes are among the largest and most complex cellular structures. Strictly speaking, membranes are aggregates, not polymers. However, the association of lipid molecules with each other creates structures that exhibit properties not shown by individual component molecules. Their insolubility in water and the flexibility of lipid aggregates give biological membranes many of their characteristics.



▲ **Figure 1.11**
Model of a membrane lipid. The molecule consists of a polar head (blue) and a nonpolar tail (yellow).

Hydrophobic interactions are discussed in Chapter 2.

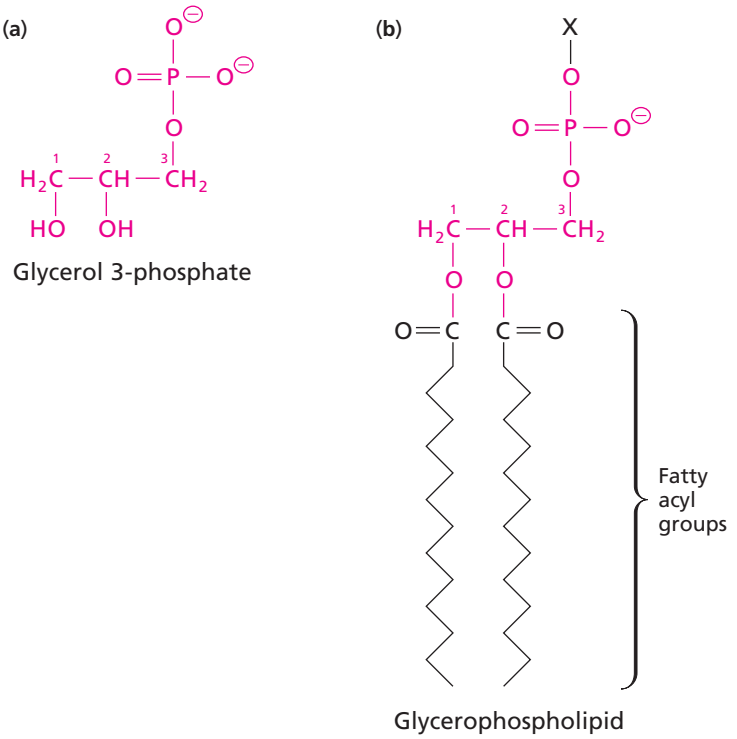


Figure 1.12 ►
Structures of glycerol 3-phosphate and a glycerophospholipid. (a) The phosphate group of glycerol 3-phosphate is polar. **(b)** In a glycerophospholipid, two nonpolar fatty acid chains are bound to glycerol 3-phosphate through ester linkages. X represents a substituent of the phosphate group.

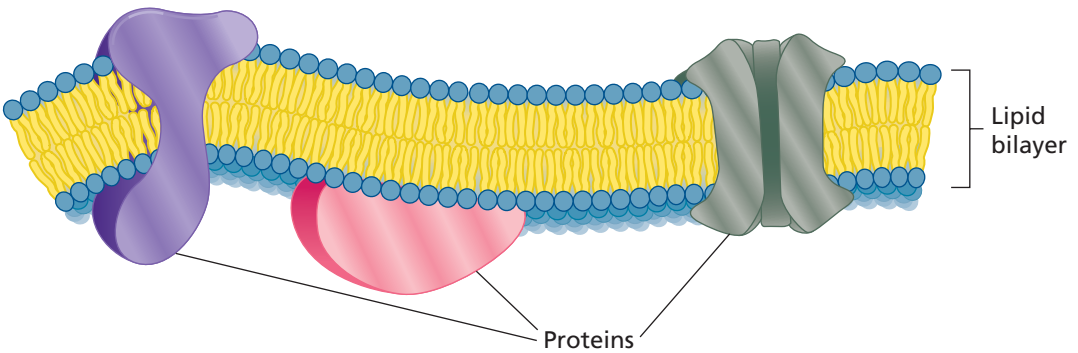
KEY CONCEPT

Most of the energy required for life is supplied by light from the sun.

Biological membranes also contain proteins as shown in Figure 1.13. Some of these membrane proteins serve as channels for the entry of nutrients and the exit of wastes. Other proteins catalyze reactions that occur specifically at the membrane surface. They are the sites of many important biochemical reactions. We will discuss lipids and biological membranes in greater detail in Chapter 9.

1.4 The Energetics of Life

The activities of living organisms do not depend solely on the biomolecules described in the preceding section and on the multitude of smaller molecules and ions found in cells. Life also requires the input of energy. Living organisms are constantly transforming energy into useful work to sustain themselves, to grow, and to reproduce. Almost all this energy is ultimately supplied by the sun.



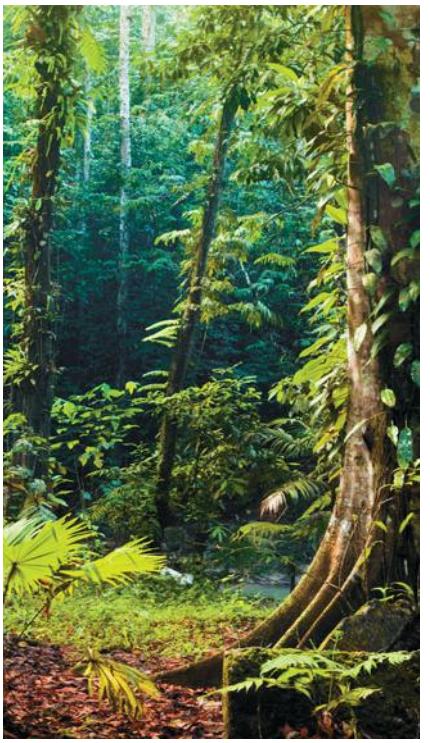
▲ **Figure 1.13**
General structure of a biological membrane. Biological membranes consist of a lipid bilayer with associated proteins. The hydrophobic tails of individual lipid molecules associate to form the core of the membrane. The hydrophilic heads are in contact with the aqueous medium on either side of the membrane. Most membrane proteins span the lipid bilayer; others are attached to the membrane surface in various ways.

Sunlight is captured by plants, algae, and photosynthetic bacteria and used for the synthesis of biological compounds. Photosynthetic organisms can be ingested as food and their component molecules used by organisms such as protozoa, fungi, nonphotosynthetic bacteria, and animals. These organisms cannot directly convert sunlight into useful biochemical energy. The breakdown of organic compounds in both photosynthetic and nonphotosynthetic organisms releases energy that can be used to drive the synthesis of new molecules and macromolecules.

Photosynthesis is one of the key biochemical processes that are essential for life, even though many species, including animals, benefit only indirectly. One of the byproducts of photosynthesis is oxygen. It is likely that Earth's atmosphere was transformed by oxygen-producing photosynthetic bacteria during the first several billion years of its history (a natural example of terraforming). In Chapter 15, we will discuss the amazing set of reactions that capture sunlight and use it to synthesize biopolymers.

The term *metabolism* describes the myriad reactions in which organic compounds are synthesized and degraded and useful energy is extracted, stored, and used. The study of the changes in energy during metabolic reactions is called *bioenergetics*. Bioenergetics is part of the field of thermodynamics, a branch of physical science that deals with energy changes. Biochemists have discovered that the basic thermodynamic principles that apply to energy flow in nonliving systems also apply to the chemistry of life.

Thermodynamics is a complex and highly sophisticated subject but we don't need to master all of its complexities and subtleties in order to understand how it can contribute to an understanding of biochemistry. We will avoid some of the complications of thermodynamics in this book and concentrate instead on using it to describe some biochemical principles (discussed in Chapter 10).



▲ **Sunlight on a tropical rain forest.** Plants convert sunlight and inorganic nutrients into organic compounds.

A. Reaction Rates and Equilibria

The rate, or speed, of a chemical reaction depends on the concentration of the reactants. Consider a simple chemical reaction where molecule A collides with molecule B and undergoes a reaction that produces products C and D.

$$A + B \longrightarrow C + D \qquad \textbf{(1.2)}$$

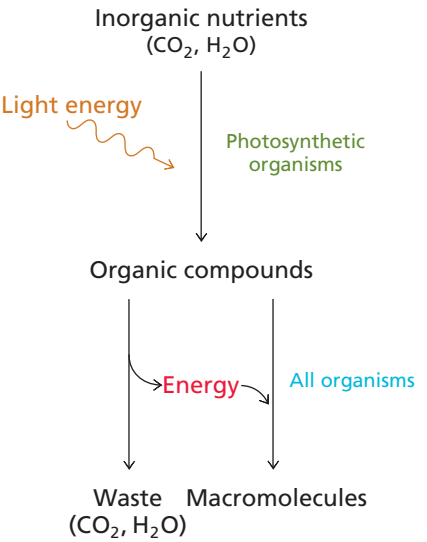
The rate of this reaction is determined by the concentrations of A and B. At high concentrations, these reactants are more likely to collide with each other; at low concentrations, the reaction might take a long time. We indicate the concentration of a reacting molecule by enclosing its symbol in square brackets. Thus, [A] means "the concentration of A"—usually expressed in moles per liter (M). The rate of the reaction is directly proportional to the product of the concentrations of A and B. This rate can be described by a proportionality constant, k, that is more commonly called a rate constant.

$$\text{rate} \propto [A][B] \qquad \text{rate} = k[A][B] \qquad \textbf{(1.3)}$$

Almost all biochemical reactions are reversible. This means that C and D can collide and undergo a chemical reaction to produce A and B. The rate of the reverse reaction will depend on the concentrations of C and D and that rate can be described by a different rate constant. By convention, the forward rate constant is k_1 and the reverse rate constant is k_{-1}. Reaction 1.4 is a more accurate way of depicting the reaction shown in Reaction 1.2.

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C + D \qquad \textbf{(1.4)}$$

If we begin a test tube reaction by mixing high concentrations of A and B, then the initial concentrations of C and D will be zero and the reaction will only proceed from left to right. The rate of the initial reaction will depend on the beginning concentrations of A and B and the rate constant k_1. As the reaction proceeds, the amount of A and B will decrease and the amount of C and D will increase. The reverse reaction will start to become significant as the products accumulate. The speed of the reverse reaction will depend on the concentrations of C and D and the rate constant k_{-1}.



▲ **Energy flow.** Photosynthetic organisms capture the energy of sunlight and use it to synthesize organic compounds. The breakdown of these compounds in both photosynthetic and nonphotosynthetic organisms generates energy needed for the synthesis of macromolecules and for other cellular requirements.

KEY CONCEPT

The rate of a chemical reaction depends on the concentrations of the reactants. The higher the concentration, the faster the reaction.

KEY CONCEPT

Almost all biochemical reactions are reversible. When the forward and reverse reactions are equal, the reaction is at equilibrium.

At some point, the rates of the forward and reverse reactions will be equal and there will be no further change in the concentrations of A, B, C, and D. In other words, the reaction will have reached equilibrium. At equilibrium,

$$k_1[A][B] = k_{-1}[C][D] \quad (1.5)$$

In many cases we are interested in the final concentrations of the reactants and products once the reaction has reached equilibrium. The ratio of product concentrations to reactant concentrations defines the equilibrium constant, K_{eq}. The equilibrium constant is also equal to the ratio of the forward and reverse rate constants and since k_1 and k_{-1} are constants, so is K_{eq}. Rearranging Equation 1.5 gives,

$$\frac{k_1}{k_{-1}} = \frac{[C][D]}{[A][B]} = K_{eq} \quad (1.6)$$

In theory, the concentrations of products and reactants could be identical once the reaction reaches equilibrium. In that case, $K_{eq} = 1$ and the forward and reverse rate constants have the same values. In most cases the value of the equilibrium constant ranges from 10^{-3} to 10^{3} meaning that the rate of one of the reactions is much faster than the other. If $K_{eq} = 10^{3}$ then the reaction will proceed mostly to the right and the final concentrations of C and D will be much higher than the concentrations of A and B. In this case, the forward rate constant (k_1) will be 1000 times greater than the reverse rate constant (k_{-1}). This means that collisions between C and D are much less likely to produce a chemical reaction than collisions between A and B.

B. Thermodynamics

If we know the energy changes associated with a reaction or process, we can predict the equilibrium concentrations. We can also predict the direction of a reaction provided we know the initial concentrations of reactants and products. The thermodynamic quantity that provides this information is the Gibbs free energy (G), named after J. Willard Gibbs who first described this quantity in 1878.

▲ **Josiah Willard Gibbs (1839–1903).** Gibbs was one of the greatest American scientists of the 19th century. He founded the modern field of chemical thermodynamics.

It turns out that molecules in solution have a certain energy that depends on temperature, pressure, concentration, and other states. The Gibbs free energy change (ΔG) for a reaction is the difference between the free energy of the products and the free energy of the reactants. The overall Gibbs free energy change has two components known as the enthalpy change (ΔH, the change in heat content) and the entropy change (ΔS, the change in randomness). A biochemical process may generate heat or absorb it from the surroundings. Similarly, a process may occur with an increase or a decrease in the degree of disorder, or randomness, of the reactants.

Starting with an initial solution of reactants and products, if the reaction proceeds to produce more products, then ΔG must be less than zero ($\Delta G < 0$). In chemistry terms, we say that the reaction is spontaneous and energy is released. When ΔG is greater than zero ($\Delta G > 0$), the reaction requires external energy to proceed and it will not yield more products. In fact, more reactants will accumulate as the reverse reaction is favored. When ΔG equals zero ($\Delta G = 0$), the reaction is at equilibrium; the rates of the forward and reverse reactions are identical and the concentrations of the products and reactants no longer change.

We are mostly interested the overall Gibbs free energy change, expressed as

$$\Delta G = \Delta H - T\Delta S \quad (1.7)$$

KEY CONCEPT

The Gibbs free energy change (ΔG) is the difference between the free energy of the products of a reaction and that of the reactants (substrates).

where T is the temperature in Kelvin.

A series of linked processes, such as the reactions of a metabolic pathway in a cell, usually proceeds only when associated with an overall negative Gibbs free energy change. Biochemical reactions or processes are more likely to occur, both to a greater extent and more rapidly, when they are associated with an increase in entropy and a decrease in enthalpy.

If we knew the Gibbs free energy of every product and every reactant, it would be a simple matter to calculate the Gibbs free energy change for a reaction by using Equation 1.8.

$$\Delta G_{\text{reaction}} = \Delta G_{\text{products}} - \Delta G_{\text{reactants}} \quad \textbf{(1.8)}$$

Unfortunately, we don't often know the absolute Gibbs free energies of every biochemical molecule. What we do know are the thermodynamic parameters associated with the *synthesis* of these molecules from simple precursors. For example, glucose can be formed from water and carbon dioxide. We don't need to know the absolute values of the Gibbs free energy of water and carbon dioxide in order to calculate the amount of enthalpy and entropy that are required to bring them together to make glucose. In fact, the heat released by the reverse reaction (breakdown of glucose to carbon dioxide and water) can be measured using a calorimeter. This gives us a value for the change in enthalpy of synthesis of glucose (ΔH). The entropy change (ΔS) for this reaction can also be determined. We can use these quantities to determine the Gibbs free energy of the reaction. The true Gibbs free energy of formation $\Delta_f G$ is the difference between the absolute free energy of glucose and that of the elements carbon, oxygen and hydrogen.

There are tables giving these Gibbs free energy values for the formation of most biological molecules. They can be used to calculate the Gibbs free energy change for a reaction in the same way that we might use absolute values as in Equation 1.9.

$$\Delta G_{\text{reaction}} = \Delta_f G_{\text{products}} - \Delta_f G_{\text{reactants}} \quad \textbf{(1.9)}$$

In this textbook we will often refer to the $\Delta_f G$ value as the Gibbs free energy of a compound since it can be easily used in calculations as though it were an absolute value. It can also be called just "Gibbs energy" by dropping the word "free."

There's an additional complication that hasn't been mentioned. For any reaction, including the degradation of glucose, the actual free energy change depends on the concentrations of reactants and products. Let's consider the hypothetical reaction in Equation 1.2. If we begin with a certain amount of A and B and none of the products C and D, then it's obvious that the reaction can only go in one direction, at least initially. In thermodynamic terms, $\Delta G_{\text{reaction}}$ is favorable under these conditions. The higher the concentrations of A and B, the more likely the reaction will occur. This is an important point that we will return to many times as we learn about biochemistry—the actual Gibbs free energy change in a reaction depends on the concentrations of the reactants and products.

What we need are some standard values of ΔG that can be adjusted for concentration. These standard values are the Gibbs free energy changes measured under certain conditions. By convention, the standard conditions are 25°C (298 K), 1 atm standard pressure, and 1.0 M concentration of all products and reactants. In most biochemical reactions, the concentration of $H^{\oplus}$ is important, and this is indicated by the pH, as will be described in the next chapter. The standard condition for biochemistry reactions is pH = 7.0, which corresponds to 10^{-7} M $H^{\oplus}$ (rather than 1.0 M as for other reactants and products). The Gibbs free energy change under these standard conditions is indicated by the symbol $\Delta G^{\circ\prime}$.

The actual Gibbs free energy is related to its standard free energy by

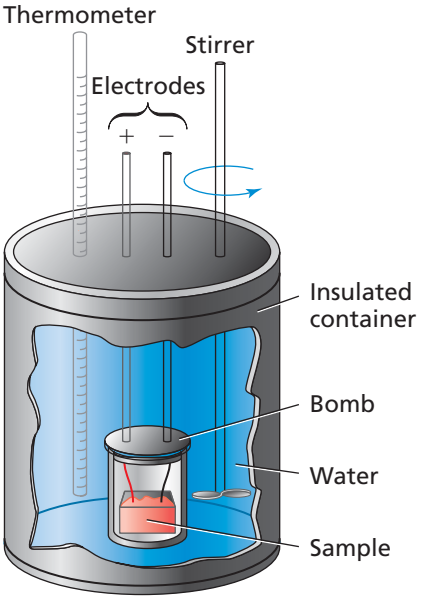
$$\Delta G_{\text{A}} = \Delta G^{\circ\prime}_{\text{A}} + RT \ln[\text{A}] \quad \textbf{(1.10)}$$

where R is the universal gas constant ($8.315\ \text{kJ}^{-1}\ \text{mol}^{-1}$) and T is the temperature in Kelvin. Gibbs free energy is expressed in units of kJ mol^{-1}. (An older unit is kcal mol^{-1}, which equals 4.184 kJ mol^{-1}.) The term $RT \ln[\text{A}]$ is sometimes given as $2.303\ RT \log[\text{A}]$.

C. Equilibrium Constants and Standard Gibbs Free Energy Changes

For a given reaction, such as that in Reaction 1.2, the actual Gibbs free energy change is related to the standard free energy change by

$$\Delta G_{\text{reaction}} = \Delta G^{\circ\prime}_{\text{reaction}} + RT \ln \frac{[\text{C}][\text{D}]}{[\text{A}][\text{B}]} \quad \textbf{(1.11)}$$



▲ The heat given off during a reaction can be determined by carrying out the reaction in a sensitive calorimeter.

The importance of the relationship between ΔG and concentration is explained in Section 10.5.

KEY CONCEPT

The standard Gibbs free energy change ($\Delta G^{\circ\prime}$) tells us the direction of a reaction when the concentrations of all products and reactants are at 1 M concentration. These conditions will never occur in living cells. Biochemists are only interested in actual Gibbs free energy changes (ΔG), which are usually close to zero. The standard Gibbs free energy change ($\Delta G^{\circ\prime}$) tells us the relative concentrations of reactants and products when the reaction reaches equilibrium.

KEY CONCEPT

$$\Delta G = \Delta G^{\circ\prime} + RT \ln \frac{[C][D]}{[A][B]}$$

at equilibrium $\Delta G^{\circ\prime} + RT \ln K_{eq} = 0$

If the reaction has reached equilibrium, the ratio of concentrations in the last term of Equation 1.11 is, by definition, the equilibrium constant (K_{eq}). When the reaction is at equilibrium there is no net change in the concentrations of reactants and products, so the actual Gibbs free energy change is zero ($\Delta G_{reaction} = 0$). This allows us to write an equation relating the standard Gibbs free energy change and the equilibrium constant. Thus, at equilibrium,

$$\Delta G^{\circ\prime}_{reaction} = -RT \ln K_{eq} = -2.303\ RT \log K_{eq} \qquad \textbf{(1.12)}$$

This important equation relates thermodynamics and reaction equilibria. Note that it is the equilibrium constant that is related to the Gibbs free energy change and not the individual rate constants described in Equations 1.6 and 1.7. It is the *ratio* of those individual rate constants that is important and not their absolute values. The forward and reverse rates might both be very slow or very fast and still give the same ratio.

D. Gibbs Free Energy and Reaction Rates

KEY CONCEPT

The rate of a reaction is not determined by the Gibbs free energy change.

Thermodynamic considerations can tell us if a reaction is favored but do not tell how quickly a reaction will occur. We know, for example, that iron rusts and copper turns green, but these reactions may take only a few seconds or many years. That's because, the rate of a reaction depends on other factors, such as the activation energy.

Activation energies are usually depicted as a hump, or barrier, in diagrams that show the progress of a reaction from left to right. In Figure 1.14, we plot the Gibbs free energy at different stages of a reaction as it goes from reactants to products. This progress is called the reaction coordinate.

The overall change in free energy (ΔG) can be negative, as shown on the left, or positive, as shown on the right. In either case, there's an excess of energy required in order for the reaction to proceed. The difference between the top of the energy peak and the energy of the product or reactant with the highest Gibbs free energy is known as the activation energy ($\Delta G‡$).

The *rate* of this reaction depends on the nature of the reaction. Using our example from Equation 1.2, if every collision between A and B is effective, then the rate is likely to be fast. On the other hand, if the orientation of individual molecules has to be exactly right for a reaction to occur then many collisions will be nonproductive and the rate will be slower. In addition to orientation, the rate depends on the kinetic energy of the individual molecules. At any given temperature some will be moving slowly when they collide and they will not have enough energy to react. Others will be moving rapidly and will carry a lot of kinetic energy.

The activation energy is meant to reflect these parameters. It is a measure of the probability that a reaction will occur. The activation energy depends on the temperature—it is lower at higher temperatures. It also depends on the concentration of reactants—at high concentrations there will be more collisions and the rate of the reaction will be faster.

The important point is that the rate of a reaction is not predictable from the overall Gibbs free energy change. Some reactions, such as the oxidation of iron or copper, will proceed very slowly because their activation energies are high.

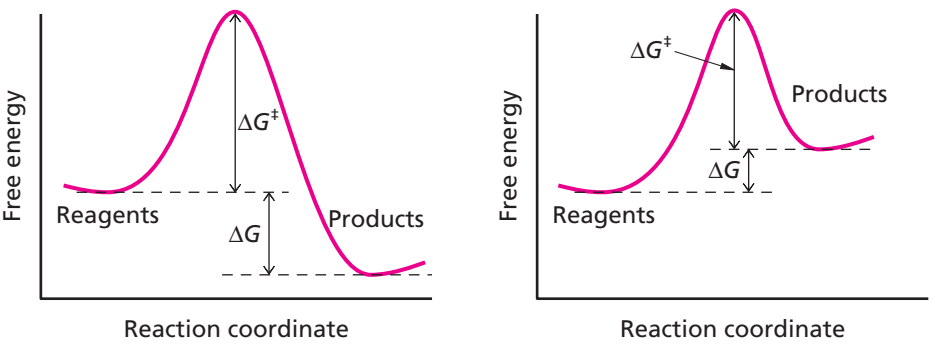


Figure 1.14 ▶
The progress of a reaction is depicted from left (reactants) to right (products). In the first diagram, the overall Gibbs free energy change is negative since the Gibbs free energy of the products is lower than that of the reactants. In order for the reaction to proceed, the reactants have to overcome an activation energy barrier ($\Delta G‡$). In the second diagram, the overall Gibbs free energy change for the reaction is positive and the minimum activation energy is smaller. This means that the reverse reaction will proceed faster than the forward reaction.

Most of the reactions that take place inside a cell are very slow in the test tube even though they are thermodynamically favored. Inside a cell the rates of the normally slow reactions are accelerated by enzymes. The rates of enzyme-catalyzed reactions can be 10^{20} times greater than the rates of the corresponding uncatalyzed reactions. We will spend some time describing how enzymes work—it is one of the most fascinating topics in biochemistry.

1.5 Biochemistry and Evolution

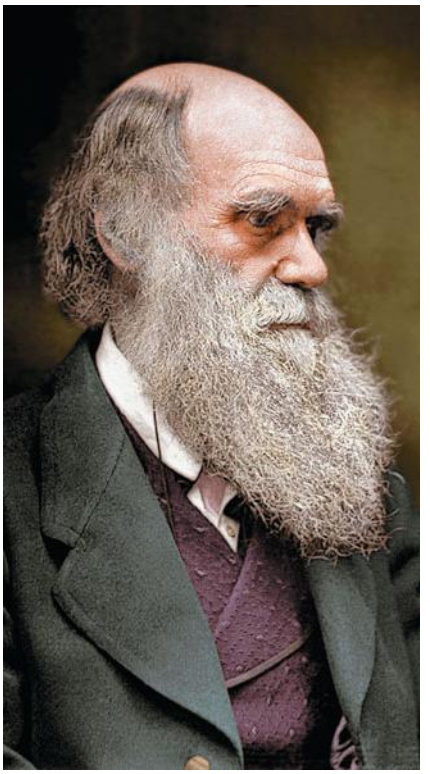
A famous geneticist, Theodosius Dobzhansky, once said, "Nothing in biology makes sense except in the light of evolution." This is also true of biochemistry. Biochemists and molecular biologists have made major contributions to our understanding of evolution at the molecular level and the evidence they have uncovered confirms and extends the data from comparative anatomy, population genetics, and paleontology. We've come a long way from the original evidence of evolution first summarized by Charles Darwin in the middle of the 19th century.

We now have a very reliable outline of the history of life and the relationships of the many diverse species in existence today. The first organisms were single cells that we would probably classify today as prokaryotes. Prokaryotes, or bacteria, do not have a membrane-bounded nucleus. Fossils of primitive bacteria-like organisms have been found in geological formations that are at least 3 billion years old. The modern species of bacteria belong to such diverse groups as the cyanobacteria, which are capable of photosynthesis, and the thermophiles, which inhabit hostile environments such as thermal hot springs.

Eukaryotes have cells that possess complex internal architecture, including a prominent nucleus. In general, eukaryotic cells are more complex and much larger than prokaryotic cells. A typical eukaryotic tissue cell has a diameter of about 25 μm (25,000 nm), whereas prokaryotic cells are typically about 1/10 that size. However, evolution has produced tremendous diversity and extreme deviations from typical sizes are common. For example, some eukaryotic unicellular organisms are large enough to be visible to the naked eye and some nerve cells in the spinal columns of vertebrates can be several feet long. There are also megabacteria that are larger than most eukaryotic cells.

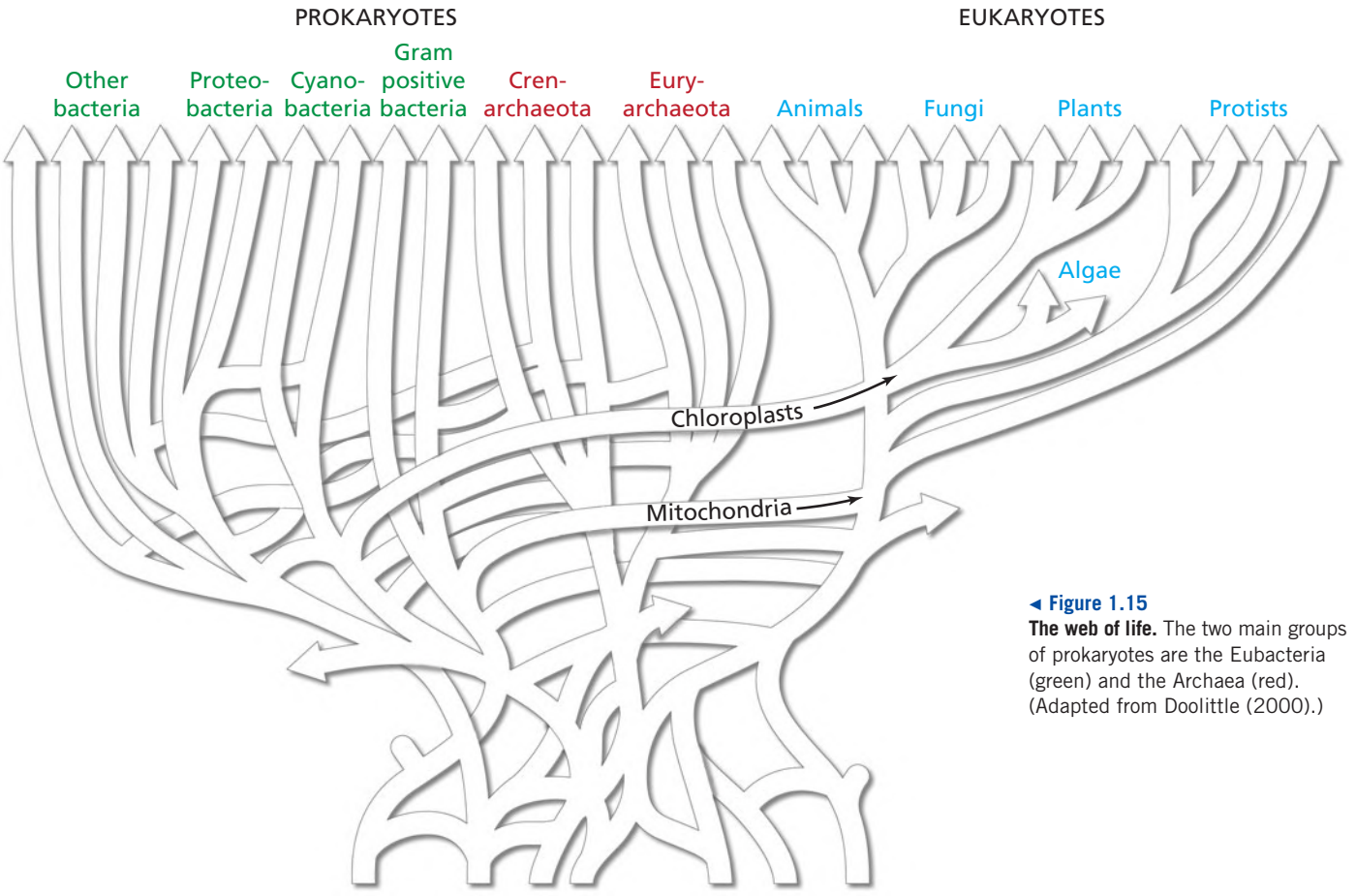
All cells on Earth (prokaryotes and eukaryotes) appear to have evolved from a common ancestor that existed more than 3 billion years ago. The evidence for common ancestry includes the presence in all living organisms of common biochemical building blocks, the same general patterns of metabolism, and a common genetic code (with rare, slight variations). We will see many examples of this evidence throughout this book. The basic plan of the primitive cell has been elaborated on with spectacular inventiveness through billions of years of evolution.

The importance of evolution for a thorough understanding of biochemistry cannot be overestimated. We will encounter many pathways and processes that only make sense



▲ **Charles Darwin (1809–1882).** Darwin published *The Origin of Species* in 1859. His theory of evolution by natural selection explains adaptive evolution.

◀ **Burgess Shale animals.** Many transitional fossils support the basic history of life that has been worked out over the past few centuries. *Pikia,* (left) is a primitive chordate from the time of the Cambrian explosion about 530 million years ago. These primitive chordates are the ancestors of all modern chordates, including humans. On the right is *Opabinia,* a primitive invertebrate.



◀ **Figure 1.15**
The web of life. The two main groups of prokaryotes are the Eubacteria (green) and the Archaea (red). (Adapted from Doolittle (2000).)

when we appreciate that they have evolved from more primitive precursors. The evidence for evolution at the molecular level is preserved in the sequences of the genes and proteins that we will study as we learn about biochemistry. In order to fully understand the fundamental principles of biochemistry we will need to examine pathways and processes in a variety of different species including bacteria and a host of eukaryotic model organisms such as yeast, fruit flies, flowering plants, mice, and humans. The importance of comparative biochemistry has been recognized for over 100 years but its value has increased enormously in the last decade with the publication of complete genome sequences. We are now able to compare the complete biochemical pathways of many different species.

The relationship of the earliest forms of life can be determined by comparing the sequences of genes and proteins in modern species. The latest evidence shows that the early forms of unicellular life exchanged genes frequently giving rise to a complicated network of genetic relationships. Eventually, the various lineages of bacteria and archaebacteria emerged, along with primitive eukaryotes. Further evolution of eukaryotes occurred when they formed a symbiotic union with bacteria, giving rise to mitochondria and chloroplasts.

The new "web of life" view of evolution (Figure 1.15) replaces a more traditional view that separated prokaryotes into two entirely separate *domains* called Eubacteria and Archaea. That distinction is not supported by the data from hundreds of sequenced genomes so we now see prokaryotes as a single large group with many diverse subgroups, some of which are shown in the figure. It is also clear that eukaryotes contain many genes that are more closely related to the old eubacterial groups as well as a minority of genes that are closer to the old achaeal groups. The early history of life seems to be dominated by rampant gene exchange between species and this has led to a web of life rather than a tree of life.

Many students are interested in human biochemistry, particularly those aspects of biochemistry that relate to health and disease. That is an exciting part of biochemistry but in order to obtain a deep understanding of who we are, we need to know where we came from. An evolutionary perspective helps explain why we can't make some vitamins

and amino acids and why we have different blood types and different tolerances for milk products. Evolution also explains the unique physiology of animals, which have adapted to using other organisms as a source of metabolic fuel.

1.6 The Cell Is the Basic Unit of Life

Every organism is either a single cell or is composed of many cells. Cells exist in a remarkable variety of sizes and shapes but they can usually be classified as either eukaryotic or prokaryotic, although some taxonomists continue to split prokaryotes into two groups: Eubacteria and Archaea.

A simple cell can be pictured as a droplet of water surrounded by a plasma membrane. The water droplet contains dissolved and suspended material including proteins, polysaccharides, and nucleic acids. The high lipid content of membranes makes them flexible and self-sealing. Membranes present impermeable barriers to large molecules and charged species. This property of membranes allows for much higher concentrations of biomolecules within cells than in the surrounding medium.

The material enclosed by the plasma membrane of a cell is called the cytoplasm. The cytoplasm may contain large macromolecular structures and subcellular membrane-bound organelles. The aqueous portion of the cytoplasm minus the subcellular structures is called the cytosol. Eukaryotic cells contain a nucleus and other internal membrane-bound organelles within the cytoplasm.

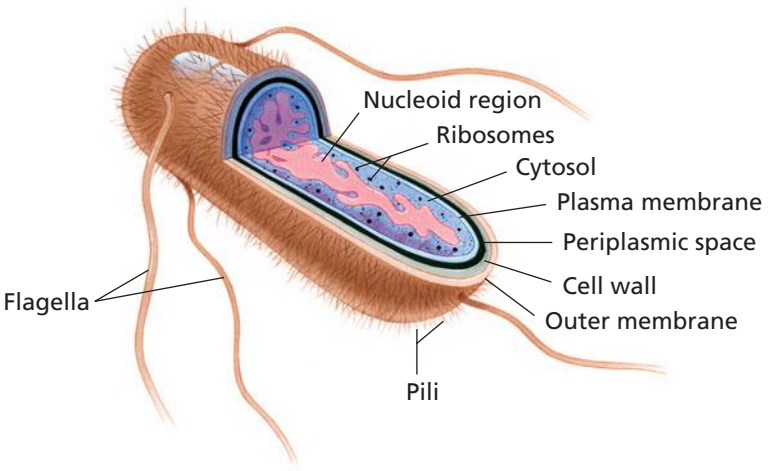
Viruses are subcellular infectious particles. They consist of a nucleic acid molecule surrounded by a protein coat and, in some cases, a membrane. Virus nucleic acid can contain as few as three genes or as many as several hundred. Despite their biological importance, viruses are not truly cells because they cannot carry out independent metabolic reactions. They propagate by hijacking the reproductive machinery of a host cell and diverting it to the formation of new viruses. In a sense, viruses are genetic parasites.

There are thousands of different viruses. Those that infect prokaryotic cells are usually called bacteriophages, or phages. Much of what we know about biochemistry is derived from the study of viruses and bacteriophages and their interaction with the cells they infect. For example, introns were first discovered in a human adenovirus like the one shown on the first page of this chapter and the detailed mapping of genes was first carried out with bacteriophage T4.

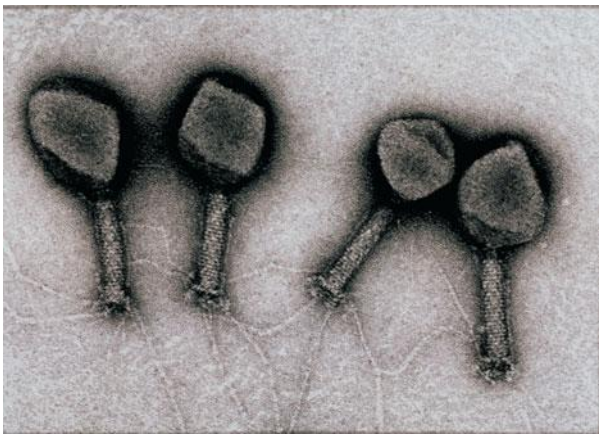
In the following two sections we will explore the structural features of typical prokaryotic and eukaryotic cells.

1.7 Prokaryotic Cells: Structural Features

Prokaryotes are usually single-celled organisms. The best studied of all living organisms is the bacterium *Escherichia coli* (Figure 1.16). This organism has served for half a century as a model biological system and many of the biochemical reactions described later in this book were first discovered in *E. coli*. *E. coli* is a fairly typical species of bacteria but some bacteria are as different from *E. coli* as we are from diatoms, daffodils and dragonflies.



◂ **Figure 1.16**
Escherichia coli. An *E. coli* cell is about 0.5 μm in diameter and 1.5 μm long. Proteinaceous fibers called flagella rotate to propel the cell. The shorter pili aid in sexual conjugation and may help *E. coli* cells adhere to surfaces. The periplasmic space is an aqueous compartment separating the plasma membrane and the outer membrane.



► **Bacteriophage T4.** Much of our current understanding of biochemistry comes from studies of bacterial viruses such as bacteriophage T4.

Much of this diversity is apparent only at the molecular level. (See Figure 1.15 for the names of some major groups of prokaryotes.)

Prokaryotes have been found in almost every conceivable environment on Earth, from hot sulfur springs to beneath the ocean floor to the insides of larger cells. They account for a significant amount of the biomass on Earth.

Prokaryotes share a number of features in spite of their differences. They lack a nucleus—their DNA is packed in a region of the cytoplasm called the nucleoid region. Many bacterial species have only 1000 genes. From a biochemist's perspective one of the most fascinating things about bacteria is that, although their chromosomes contain a relatively small number of genes, they carry out most of the fundamental biochemical reactions found in all cells, including our own. Hundreds of bacterial genomes have been completely sequenced and it is now possible to begin to define the minimum number of enzymes that are consistent with life.

Most bacteria have no internal membrane compartments, although there are many exceptions. The plasma membrane is usually surrounded by a cell wall made of a rigid network of covalently linked carbohydrate and peptide chains. This cell wall confers the characteristic shape of an individual species of bacteria. Despite its mechanical strength, the cell wall is porous. In addition to the cell wall most bacteria, including *E. coli*, possess an outer membrane consisting of lipids, proteins, and lipids linked to polysaccharides. The space between the inner plasma membrane and the outer membrane is called the periplasmic space. It is the major membrane-bound compartment in bacteria and plays a crucial role in some important biochemical processes.

Many bacteria have protein fibers, called pili, on their outer surface. The pili serve as attachment sites for cell-cell interactions. Many species have one or more flagella. These are long, whip-like structures that can be rotated like the propeller on a boat thus driving the bacterium through its aqueous environment.

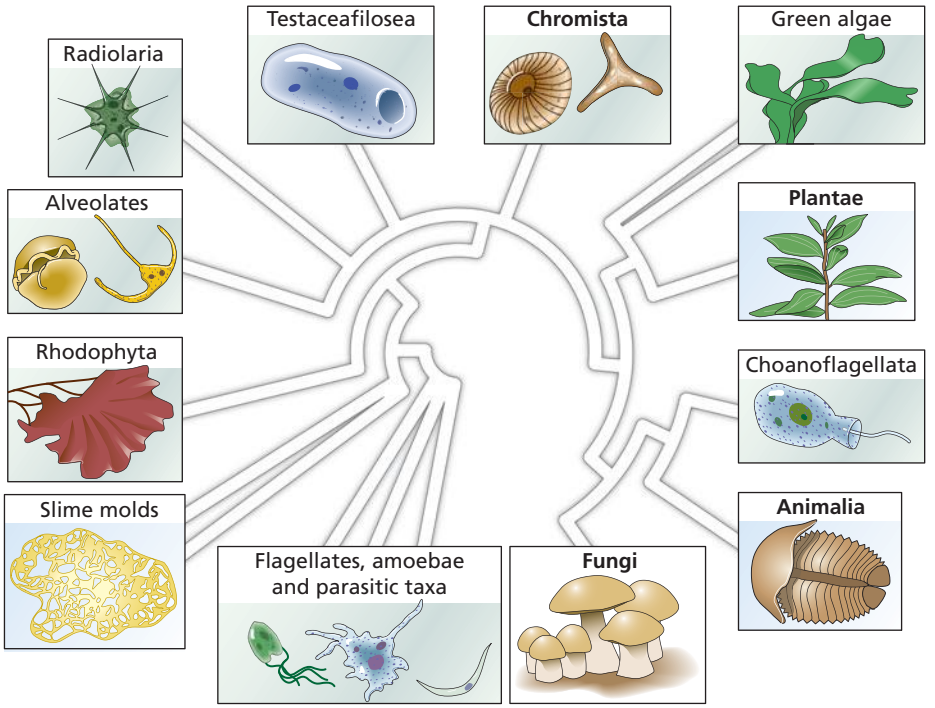
The small size of prokaryotes provides a high ratio of surface area to volume. Simple diffusion is therefore an adequate means for distributing nutrients throughout the cytoplasm. One of the prominent macromolecular structures in the cytoplasm is the ribosome—a large RNA-protein complex required for protein synthesis. All living cells have ribosomes but we will see later that bacterial ribosomes differ from eukaryotic ribosomes in significant details.

▲ **Max Delbruck and Salvatore Luria.** Max Delbruck (seated) and Salvatore Luria at the Cold Spring Harbor Laboratories in 1953. Delbruck and Luria founded the "phage group," a group of scientists who worked on the genetics and biochemistry of bacteria and bacteriophage in the 1940s, 1950s, and 1960s.

1.8 Eukaryotic Cells: Structural Features

Eukaryotes include plants, animals, fungi, and protists. Protists are mostly small, single-celled organisms that don't fit into one of the other classes. Along with bacteria these four groups make up the five kingdoms of life according to one popular classification scheme. (Older schemes retain the four eukaryotic kingdoms but divide the bacteria into Eubacteria and Archaea.)

As members of the animal kingdom we are mostly aware of other animals. As relatively large organisms we tend to focus on the large scale. Hence, we know about plants and mushrooms but not microscopic species.



◀ **Figure 1.17**
The eukaryotic tree of life. The traditional Plantae, Animalia, and Fungi kingdoms are branches within the much larger "kingdom" of Protists.

The latest trees of eukaryotes help us understand the diversity of the protist kingdom. As shown in Figure 1.17, the animal, plant, and fungal "kingdoms" occupy relatively small branches on the eukaryotic tree of life.

Eukaryotic cells are surrounded by a single plasma membrane unlike bacteria, which usually have a double membrane. The most obvious feature that distinguishes eukaryotes from prokaryotes is the presence of a membrane-bound nucleus in eukaryotes. In fact, eukaryotes are defined by the presence of a nucleus (from the Greek: *eu*-, "true" and *karuon*, "nut" or "kernel.").

As mentioned earlier, eukaryotic cells are almost always larger than bacterial cells, commonly 1000-fold greater in volume. Because of their large size complex internal structures and mechanisms are required for rapid transport and communication both inside the cell and to and from the external medium. A mesh of protein fibers called the cytoskeleton extends throughout the cell contributing to cell shape and to the management of intracellular traffic.

Almost all eukaryotic cells contain additional internal membrane-bound compartments called organelles. The specific functions of organelles are often closely tied to their physical properties and structures. Nevertheless, a significant number of specific biochemical processes occur in the cytosol and the cytosol, like organelles, is highly organized.

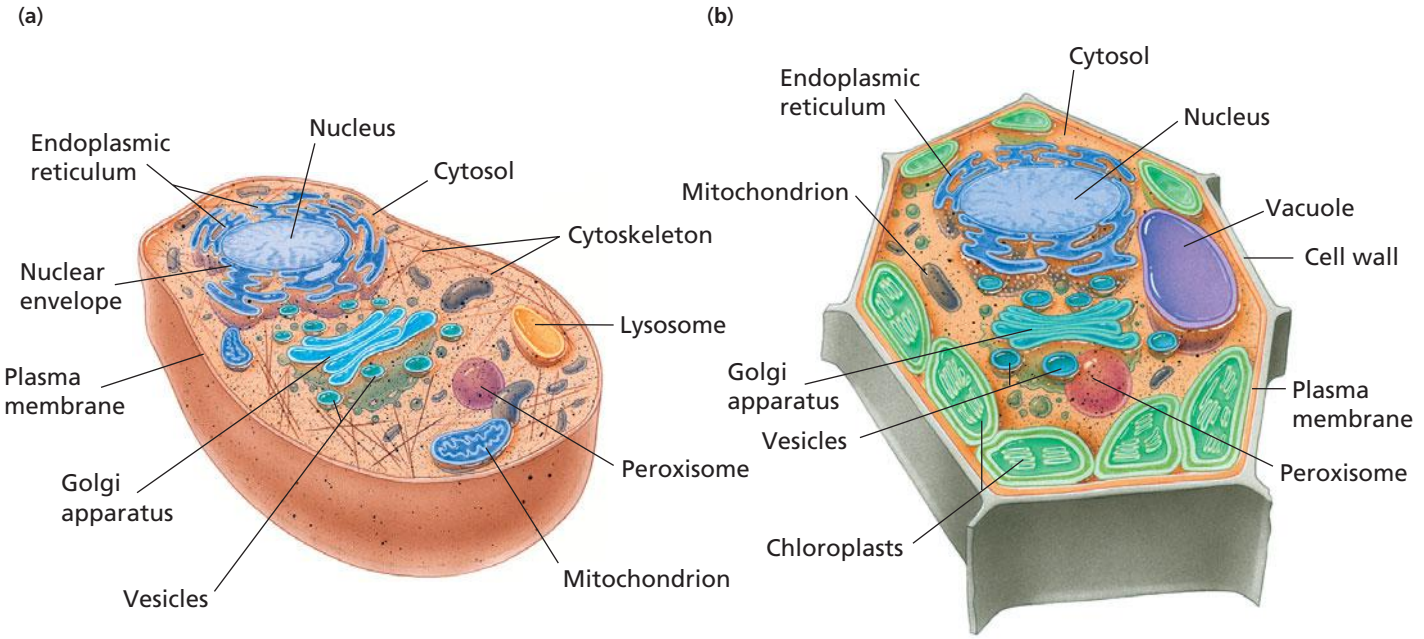
The interior of a eukaryotic cell contains an intracellular membrane network. Independent organelles, including the nucleus, mitochondria, and chloroplasts, are embedded in this membrane system that pervades the entire cell. Materials flow within paths defined by membrane walls and tubules. The intracellular traffic of materials between compartments is rapid, highly selective, and closely regulated.

Figure 1.18 on the next page shows typical animal and plant cells. Both types have a nucleus, mitochondria, and a cytoskeleton. Plant cells also contain chloroplasts and vacuoles and are often surrounded by a rigid cell wall. Chloroplasts, also found in algae and some other protists, are the sites of photosynthesis. Plant cell walls are mostly composed of cellulose, one of the polysaccharides described in Section 1.3B.

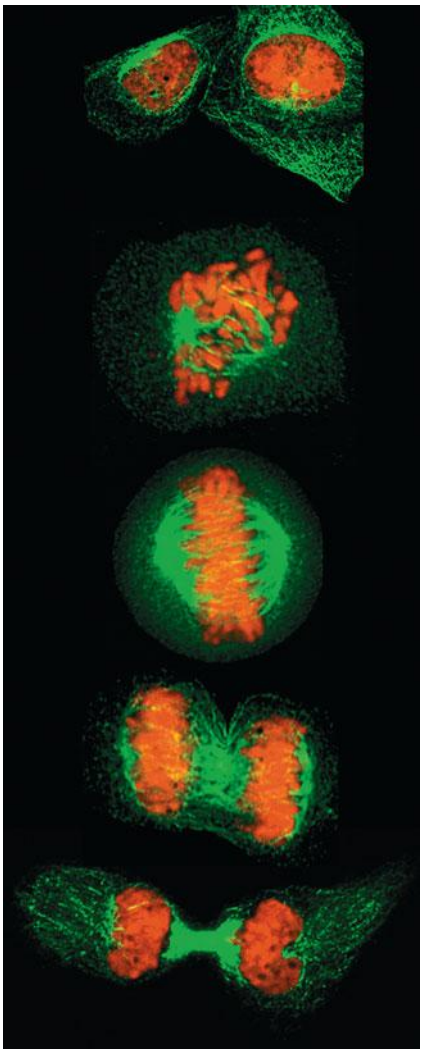
Most multicellular eukaryotes contain tissues. Groups of similarly specialized cells within tissues are surrounded by an extracellular matrix containing proteins and polysaccharides. The matrix physically supports the tissue and in some cases directs cell growth and movement.

KEY CONCEPT

Animals are a relatively small, highly specialized, branch on the tree of life.



▲ **Figure 1.18**
Eukaryotic cells. (a) Composite animal cell. Animal cells are typical eukaryotic cells containing organelles and structures also found in protists, fungi, and plants. **(b)** Composite plant cell. Most plant cells contain chloroplasts, the sites of photosynthesis in plants and algae; vacuoles, large, fluid-filled organelles containing solutes and cellular wastes; and rigid cell walls composed mostly of cellulose.



A. The Nucleus

The nucleus is usually the most obvious structure in a eukaryotic cell. It is structurally defined by the nuclear envelope, a membrane with two layers that join at protein-lined nuclear pores. The nuclear envelope is connected to the endoplasmic reticulum (see below). The nucleus is the control center of the cell containing 95% of its DNA, which is tightly packed with positively charged proteins called histones and coiled into a dense mass called chromatin. Replication of DNA and transcription of DNA into RNA occur in the nucleus. Many eukaryotes have a dense mass in the nucleus called the nucleolus. The nucleolus is a major site of RNA synthesis and the site of assembly of ribosomes.

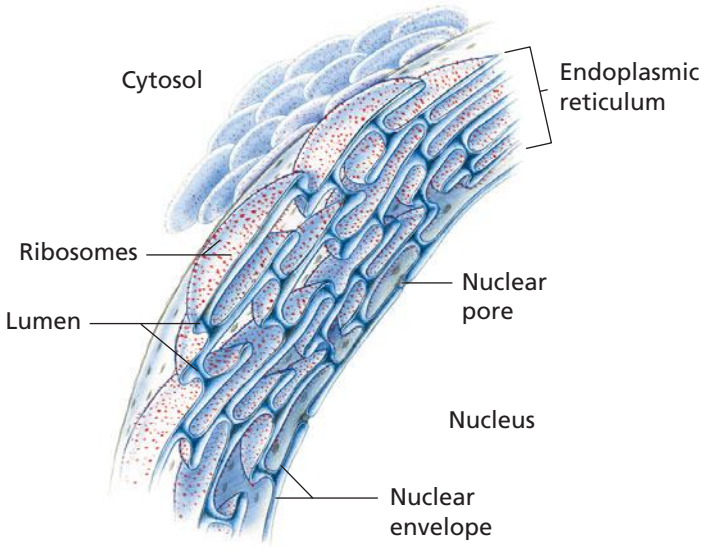
Most eukaryotes contain far more DNA than do prokaryotes. Whereas the genetic material, or genome, of prokaryotes is usually a single circular molecule of DNA, the eukaryotic genome is organized as multiple linear chromosomes. In eukaryotes new DNA and histones are synthesized in preparation for cell division and the chromosomal material condenses and separates into two identical sets of chromosomes. This process is called mitosis (Figure 1.19). The cell is then pinched in two to complete cell division.

Most eukaryotes are diploid—they contain two complete sets of chromosomes. From time to time eukaryotic cells undergo meiosis resulting in the production of four haploid cells each with a single set of chromosomes. Two haploid cells—eggs and sperm, for example—can then fuse to regenerate a typical diploid cell. This process is one of the key features of sexual reproduction in eukaryotes.

B. The Endoplasmic Reticulum and Golgi Apparatus

A network of membrane sheets and tubules called the endoplasmic reticulum (ER) extends from the outer membrane of the nucleus. The aqueous region enclosed within the endoplasmic reticulum is called the lumen. In many cells part of the surface of the endoplasmic reticulum is coated with ribosomes that are actively synthesizing proteins.

◀ **Figure 1.19**
Mitosis. The five stages of mitosis are shown. Chromosomes (red) condense and line up in the center of the cell. Spindle fibers (green) are responsible for separating the recently duplicated chromosomes.

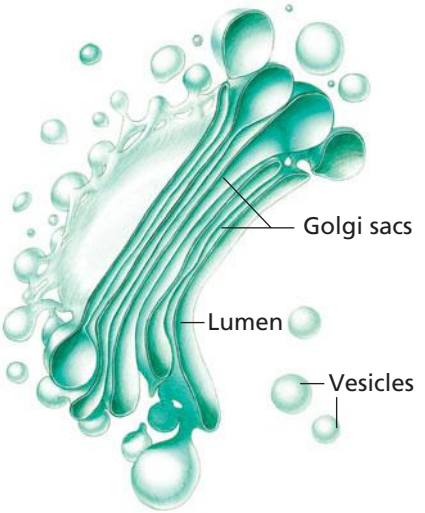


◀ **Nuclear envelope and endoplasmic reticulum (ER) of a eukaryotic cell.**

Protein synthesis, sorting, and secretion are described in Chapter 22.

As synthesis continues the protein is translocated through the membrane into the lumen. Proteins destined for export from the cell are completely extruded through the membrane into the lumen where they are packaged in membranous vesicles. These vesicles travel through the cell and fuse with the plasma membrane releasing their contents into the extracellular space. The synthesis of proteins destined to remain in the cytosol occurs at ribosomes that are not bound to the endoplasmic reticulum.

A complex of flattened, fluid-filled, membranous sacs called the Golgi apparatus is often found close to the endoplasmic reticulum and the nucleus. Vesicles that bud off from the endoplasmic reticulum fuse with the Golgi apparatus. The proteins carried by the vesicles may be chemically modified as they pass through the layers of the Golgi apparatus. The modified proteins are then sorted, packaged in new vesicles, and transported to specific destinations inside or outside the cell. The Golgi apparatus was discovered by Camillo Golgi in the 19th century (Nobel Laureate, 1906), although it wasn't until many decades later that its role in protein secretion was established.



▲ **Golgi apparatus.** The Golgi apparatus is responsible for the modification and sorting of proteins that have been transported to the Golgi apparatus by vesicles from the ER. Vesicles budding off the Golgi apparatus carry modified material to destinations inside and outside the cell.

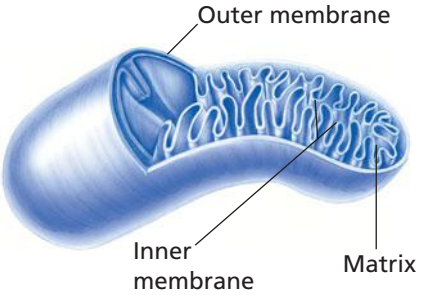
C. Mitochondria and Chloroplasts

Mitochondria and chloroplasts have central roles in energy transduction. Mitochondria are the main sites of oxidative energy metabolism. They are found in almost all eukaryotic cells. Chloroplasts are the sites of photosynthesis in plants and algae.

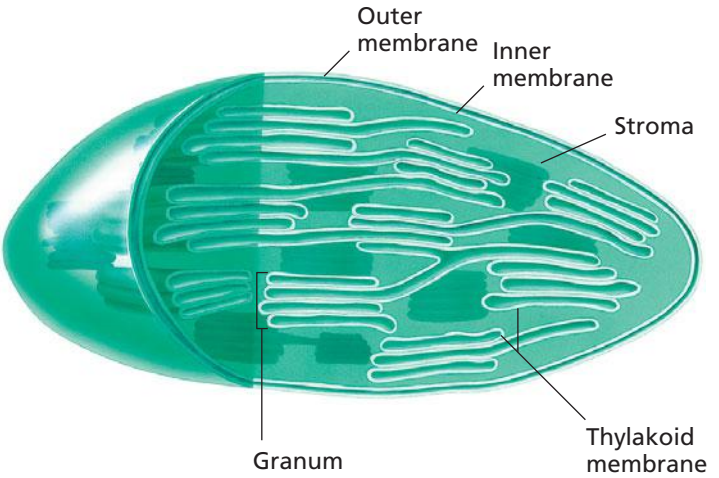
The mitochondrion has an inner and an outer membrane. The inner membrane is highly folded, resulting in a surface area three to five times that of the outer membrane. It is impermeable to ions and most metabolites. The aqueous phase enclosed by the inner membrane is called the mitochondrial matrix. Many of the enzymes involved in aerobic energy metabolism are found in the inner membrane and the matrix.

Mitochondria come in many sizes and shapes. The standard jellybean-shaped mitochondrion shown here is found in many cell types but some mitochondria are spherical or have irregular shapes.

The most important role of the mitochondrion is to oxidize organic acids, fatty acids, and amino acids to carbon dioxide and water. Much of the released energy is conserved in the form of a proton concentration gradient across the inner mitochondrial membrane. This stored energy is used to drive the conversion of adenosine diphosphate (ADP) and inorganic phosphate (P_i) to the energy-rich molecule ATP in a phosphorylation process that will be described in detail in Chapter 14. ATP is then used by the cell for such energy-requiring processes as biosynthesis, transport of certain molecules and ions against concentration and charge gradients, and generation of mechanical force for such purposes as locomotion and muscle contraction. The number of mitochondria found in cells varies widely. Some eukaryotic cells contain only a few mitochondria whereas others have thousands.



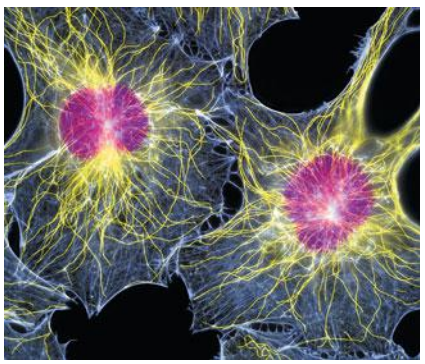
▲ **Mitochondrion.** Mitochondria are the main sites of energy transduction in aerobic eukaryotic cells. Carbohydrates, fatty acids, and amino acids are metabolized in this organelle.



► **Chloroplast.** Chloroplasts are the sites of photosynthesis in plants and algae. Light energy is captured by pigments associated with the thylakoid membrane and used to convert carbon dioxide and water to carbohydrates.

Photosynthetic plant cells contain chloroplasts as well as mitochondria. Like mitochondria, chloroplasts have an outer membrane and a complex, highly folded, inner membrane called the thylakoid membrane. Part of the inner membrane forms flattened sacs called grana (singular, granum). The thylakoid membrane, which is suspended in the aqueous stroma, contains chlorophyll and other pigments involved in the capture of light energy. Ribosomes and several circular DNA molecules are also suspended in the stroma. In chloroplasts the energy captured from light is used to drive the formation of carbohydrates from carbon dioxide and water.

Mitochondria and chloroplasts are derived from bacteria that entered into internal symbiotic relationships with primitive eukaryotic cells more than 1 billion years ago. Evidence for the endosymbiotic (*endo-*, "within") origin of mitochondria and chloroplasts includes the presence within these organelles of separate, small genomes and specific ribosomes that resemble those of bacteria. In recent years scientists have compared the sequences of mitochondrial and chloroplast genes (and proteins) with those of many species of bacteria. These studies in molecular evolution have shown that mitochondria are derived from primitive members of a particular group of bacteria called proteobacteria. Chloroplasts are descended from a distantly related class of photosynthetic bacteria called cyanobacteria.



▲ **Micrographs of fluorescently labeled actin filaments and microtubules in mammalian cells. (Left)** Actin filaments in rat muscle cells. **(Right)** Microtubules in human endothelial cells.

D. Specialized Vesicles

Eukaryotic cells contain specialized digestive vesicles called lysosomes. These vesicles are surrounded by a single membrane that encloses a highly acidic interior. The acidity is maintained by proton pumps embedded in the membrane. Lysosomes contain a variety of enzymes that catalyze the breakdown of cellular macromolecules such as proteins and nucleic acids. They can also digest large particles such as retired mitochondria and bacteria ingested by the cell. Lysosomal enzymes are much less active at the near-neutral pH of the cytosol than they are under the acidic conditions inside the lysosome. The compartmentalization of lysosomal enzymes keeps them from accidentally catalyzing the degradation of macromolecules in the cytosol.

Peroxisomes are present in all animal cells and many plant cells. Like lysosomes, they are surrounded by a single membrane. Peroxisomes carry out oxidation reactions, some of which produce the toxic compound hydrogen peroxide, (H_2O_2). Some hydrogen peroxide is used for the oxidation of other compounds. Excess hydrogen peroxide is destroyed by the action of the peroxisomal enzyme catalase, which catalyzes the conversion of hydrogen peroxide to water and oxygen.

Vacuoles are fluid-filled vesicles surrounded by a single membrane. They are common in mature plant cells and some protists. These vesicles are storage sites for water, ions, and nutrients such as glucose. Some vacuoles contain metabolic waste products and some contain enzymes that can catalyze the degradation of macromolecules no longer needed by the plant.

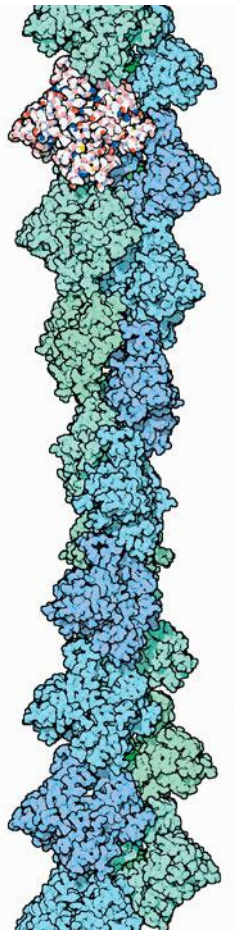
E. The Cytoskeleton

The cytoskeleton is a protein scaffold required for support, internal organization, and even movement of the cell. Some types of animal cells contain a dense cytoskeleton but it is much less prominent in most other eukaryotic cells. The cytoskeleton consists of three types of protein filaments: actin filaments, microtubules, and intermediate filaments. All three types are built of individual protein molecules that combine to form threadlike fibers.

Actin filaments (also called microfilaments) are the most abundant cytoskeletal component. They are composed of a protein called actin that forms ropelike threads with a diameter of about 7 nm. Actin has been found in all eukaryotic cells and is frequently the most abundant protein in the cell. It is also one of the most evolutionarily conserved proteins. This is evidence that actin filaments were present in the ancestral eukaryotic cell from which all modern eukaryotes are descended.

Microtubules are strong, rigid fibers frequently packed in bundles. They have a diameter of about 22 nm—much thicker than actin filaments. Microtubules are composed of a protein called tubulin. Microtubules serve as a kind of internal skeleton in the cytoplasm, but they also form the mitotic spindle during mitosis. In addition, microtubules can form structures capable of directed movement, such as cilia. The flagella that propel sperm cells are an example of very long cilia—they are not related to bacterial flagella. The waving motion of cilia is driven by energy from ATP.

Intermediate filaments are found in the cytoplasm of most eukaryotic cells. These filaments have diameters of approximately 10 nm, which makes them intermediate in size compared to actin filaments and microtubules. Intermediate filaments line the inside of the nuclear envelope and extend outward from the nucleus to the periphery of the cell. They help the cell resist external mechanical stresses.



▲ **Actin.** Actin filament showing the organization in individual subunits of the protein actin. (Courtesy David S. Goodsell)

1.9 A Picture of the Living Cell

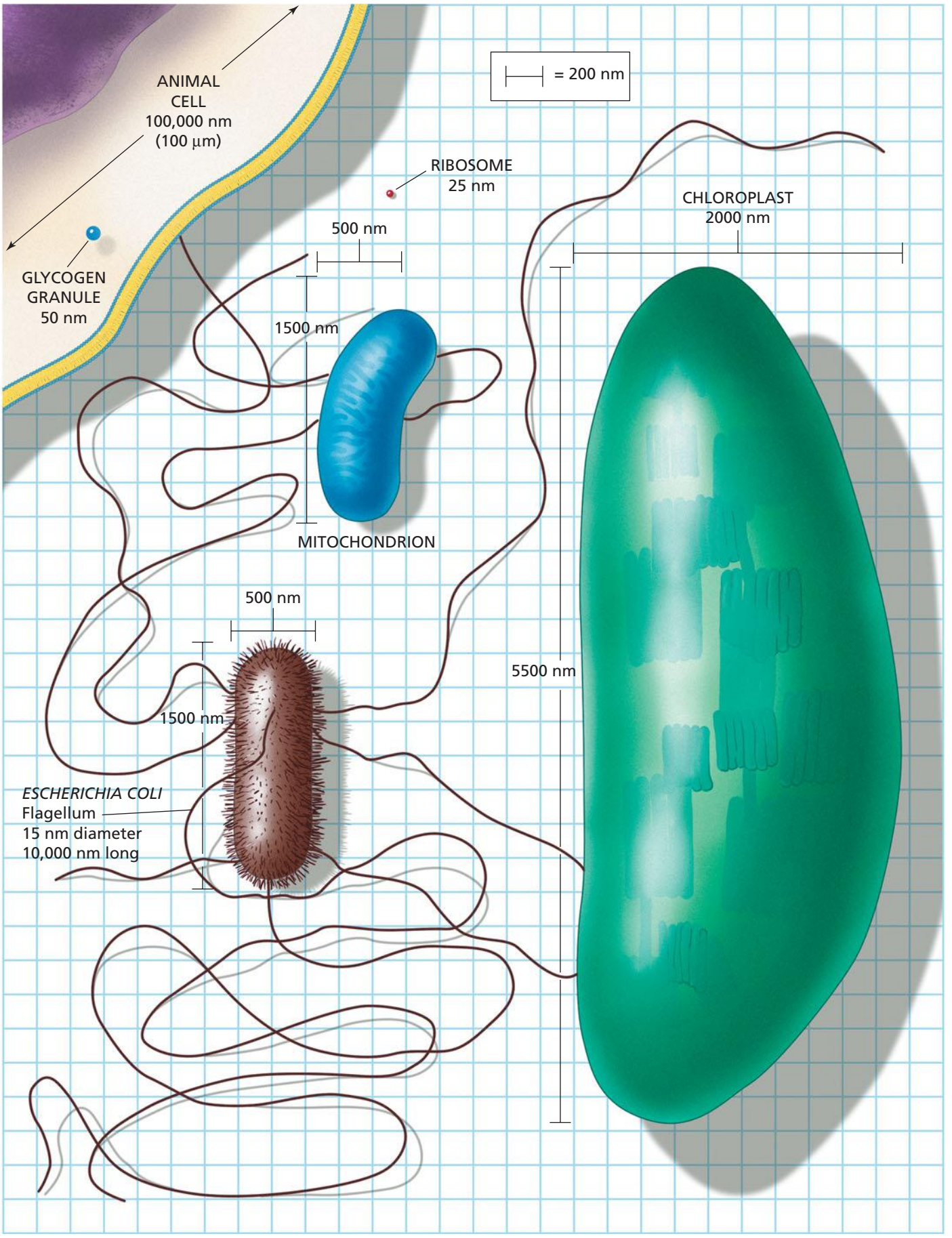
We have now introduced the major structures found within cells and described their roles. These structures are immense compared to the molecules and polymers that will be our focus for the rest of this book. Cells contain thousands of different metabolites and many millions of molecules. In the cytosol of every cell there are hundreds of different enzymes, each acting specifically on only one or possibly a few related metabolites. There may be 100,000 copies of some enzymes per cell but only a few copies of other enzymes. Each enzyme is bombarded with potential substrates.

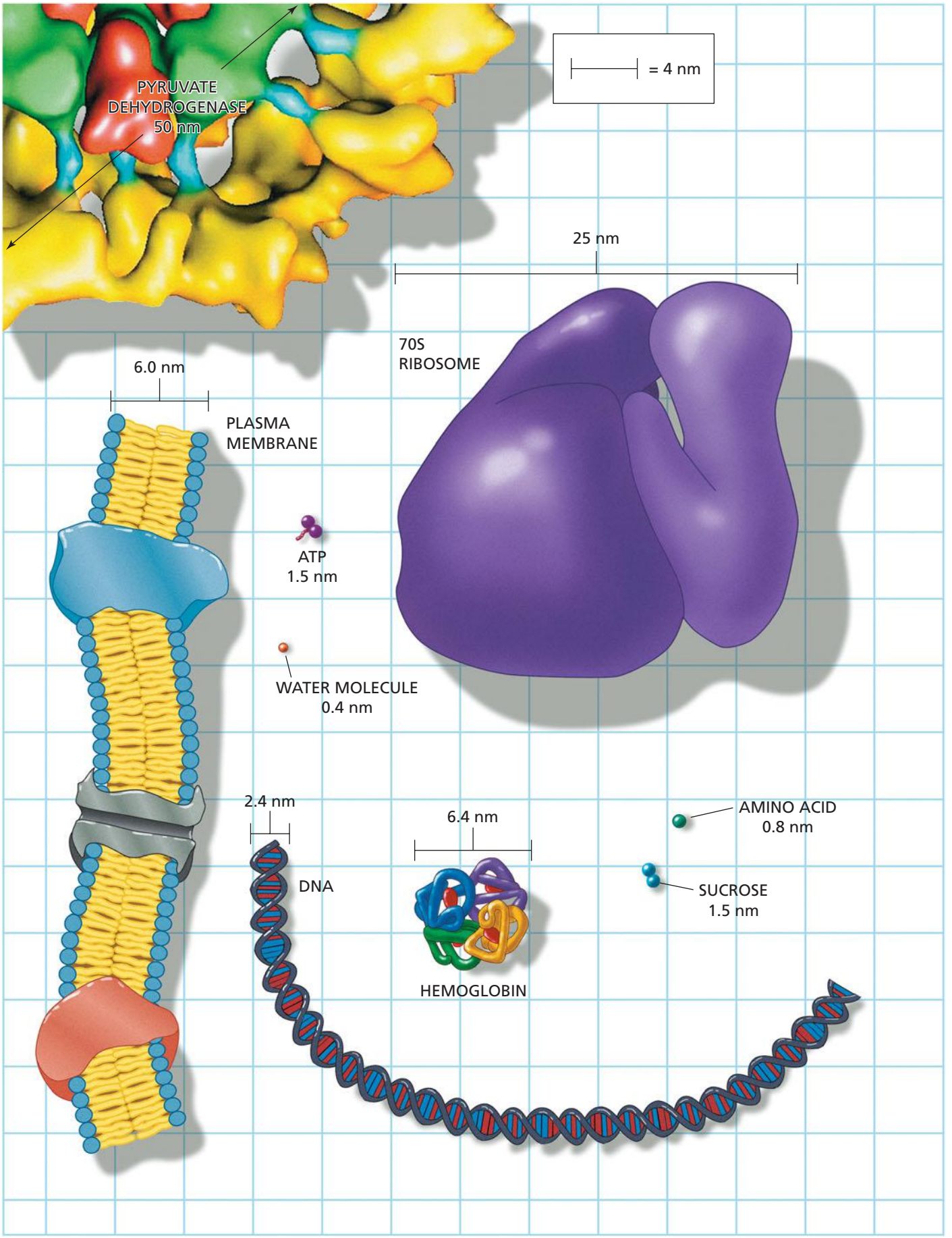
Molecular biologist and artist David S. Goodsell has produced captivating images showing the molecular contents of an *E. coli* cell magnified 1 million times (Figure 1.20 on page 26). Approximately 600 cubes of this size represent the volume of the *E. coli* cell. At this scale individual atoms are smaller than the dot in the letter *i* and small metabolites are barely visible. Proteins are the size of a grain of rice.

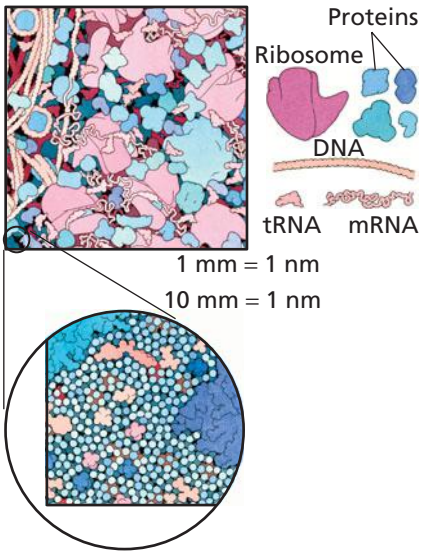
A drawing of the molecules in a cell shows how densely packed the cytoplasm can be, but it cannot give a sense of activity at the atomic scale. All the molecules in a cell are moving and colliding with each other. The collisions between molecules are fully elastic—the energy of a collision is conserved in the energy of the rebound. As molecules bounce off each other they travel a wildly crooked path in space, called the random walk of diffusion. For a small molecule such as water, the mean distance traveled between collisions is less than the dimensions of the molecule and the path includes many reversals of direction. Despite its convoluted path, a water molecule can diffuse the length of an *E. coli* cell in 1/10 second.

An enzyme and a small molecule will collide 1 million times per second. Under these conditions, a rate of catalysis typical of many enzymes could be achieved even if only 1 in about 1000 collisions results in a reaction. Nevertheless, some enzymes catalyze reactions with an efficiency far greater than 1 reaction per 1000 collisions. In fact, a few enzymes catalyze reactions with almost every molecule of substrate their active sites encounter—an example of the astounding potency of enzyme-directed chemistry. The study of the reaction rates of enzymes, or enzyme kinetics, is one of the most fundamental aspects of biochemistry. It will be covered in Chapter 6.

Lipids in membranes also diffuse vigorously, though only within the two-dimensional plane of the lipid bilayer. Lipid molecules exchange places with neighboring


= 200 nm
ANIMAL CELL
100,000 nm
(100 μm)
RIBOSOME
25 nm
CHLOROPLAST
2000 nm
GLYCOGEN GRANULE
50 nm
500 nm
1500 nm
MITOCHONDRION
500 nm
5500 nm
1500 nm
ESCHERICHIA COLI
Flagellum
15 nm diameter
10,000 nm long


PYRUVATE DEHYDROGENASE 50 nm
= 4 nm
25 nm
70S RIBOSOME
6.0 nm
PLASMA MEMBRANE
ATP 1.5 nm
WATER MOLECULE 0.4 nm
2.4 nm
DNA
6.4 nm
HEMOGLOBIN
AMINO ACID 0.8 nm
SUCROSE 1.5 nm



▲ **Figure 1.20**
Portion of the cytosol of an *E. coli* cell. The top illustration, in which the contents are magnified 1 million times, represents a window 100 x 100 nm. Proteins are in shades of blue and green. Nucleic acids are in shades of pink. The large structures are ribosomes. Water and small metabolites are not shown. The contents in the round inset are magnified 10 million times, showing water and other small molecules.

molecules in membranes about 6 million times per second. Some membrane proteins can also diffuse rapidly within the membrane.

Large molecules diffuse more slowly than small ones. In eukaryotic cells the diffusion of large molecules such as enzymes is retarded even further by the complex network of the cytoskeleton. Large molecules diffuse across a given distance as much as 10 times more slowly in the cytosol than in pure water.

The full extent of cytosolic organization is not yet known. A number of proteins and enzymes form large complexes that carry out a series of reactions. We will encounter several such complexes in our study of metabolism. They are often referred to as protein machines. This arrangement has the advantage that metabolites pass directly from one enzyme to the next without diffusing away into the cytosol. Many researchers are sympathetic to the idea that the cytosol is not merely a random mixture of soluble molecules but is highly organized in contrast to the long-held impression that simple solution chemistry governs cytosolic activity. The concept of a highly organized cytosol is a relatively new idea in biochemistry. It may lead to important new insights about how cells work at the molecular level.

1.10 Biochemistry Is Multidisciplinary

One of the goals of biochemists is to integrate a large body of knowledge into a molecular explanation of life. This has been, and continues to be, a challenging task but, in spite of the challenges, biochemists have made a great deal of progress toward defining and understanding the basic reactions common to all cells.

The discipline of biochemistry does not exist in a vacuum. We have already seen how physics, chemistry, cell biology, and evolution contribute to an understanding of biochemistry. Related disciplines, such as physiology and genetics, are also important. In fact, many scientists no longer consider themselves to be just biochemists but are also knowledgeable in several related fields.

Because all aspects of biochemistry are interrelated it is difficult to present one topic without referring to others. For example, function is intimately related to structure and the regulation of individual enzyme activities can be appreciated only in the context of a series of linked reactions. The interrelationship of biochemistry topics is a problem for both students and teachers in an introductory biochemistry course. The material must be presented in a logical and sequential manner but there is no universal sequence of topics that suits every course, or every student. Fortunately, there is general agreement on the broad outline of an approach to understanding the basic principles of biochemistry and this textbook follows that outline. We begin with an introductory chapter on water. We will then describe the structures and functions of proteins and enzymes, carbohydrates, and lipids. The third part of the book makes use of structural information to describe metabolism and its regulation. Finally, we will examine nucleic acids and the storage and transmission of biological information.

Some courses may cover the material in a slightly different order. For example, the structures of nucleic acids can be described before the metabolism section. Wherever possible, we have tried to write chapters so that they can be covered in different orders in a course depending on the particular needs and interests of the students.

Appendix The Special Terminology of Biochemistry

Most biochemical quantities are specified using Système International (SI) units. Some common SI units are listed in Table 1.1 Many biochemists still use more traditional units, although these are rapidly disappearing from the scientific literature. For example, protein chemists sometimes use the angstrom (Å) to report interatomic distances; 1 Å is equal to 0.1 nm, the preferred SI unit. Calories (cal) are sometimes used instead of joules (J); 1 cal is equal to 4.184 J.

The standard SI unit of temperature is the Kelvin, but temperature is most commonly reported in degrees Celsius (°C). One degree Celsius is equal in magnitude to 1 Kelvin, but the Celsius scale begins at the freezing point of water (0°C) and 100°C is

TABLE 1.1 SI units commonly used in biochemistry

Physical quantity	SI unit	Symbol
Length	meter	m
Mass	gram	g
Amount	mole	mol
Volume	liter[a]	L
Energy	joule	J
Electric potential	volt	V
Time	second	s
Temperature	Kelvin[b]	K

[a]1 liter = 1000 cubic centimeters.
[b]273 K = 0° C.

Table 1.2 Prefixes commonly used with SI units

Prefix	Symbol	Multiplication factor
giga-	G	10^9
mega-	M	10^6
kilo-	k	10^3
deci-	d	10^{-1}
centi-	c	10^{-2}
milli-	m	10^{-3}
micro-	μ	10^{-6}
nano-	n	10^{-9}
pico-	p	10^{-12}
femto-	f	10^{-15}

the boiling point of water at 1 atm. This scale is often referred to as the centigrade scale (*centi-* = 1/100). Absolute zero is −273 °C, which is equal to 0 K. In warm-blooded mammals biochemical reactions occur at body temperature (37°C in humans).

Very large or very small numerical values for some SI units can be indicated by an appropriate prefix. The commonly used prefixes and their symbols are listed in Table 1.2. In addition to the standard SI units employed in all fields, biochemistry has its own special terminology; for example, biochemists use convenient abbreviations for biochemicals that have long names.

The terms RNA and DNA are good examples. They are shorthand versions of the long names ribonucleic acid and deoxyribonucleic acid. Abbreviations such as these are very convenient, and learning to associate them with their corresponding chemical structures is a necessary step in mastering biochemistry. In this book, we will describe common abbreviations as each new class of compounds is introduced.

Selected Readings

Chemistry

Bruice, P. Y. (2011). *Organic Chemistry*, 6th ed. (Upper Saddle River, NJ: Prentice Hall).

Tinoco, I., Sauer, K., Wang, J. C., and Puglisi, J. D. (2002). *Physical Chemistry: Principles and Applications in Biological Sciences*, 4th ed. (Upper Saddle River, NJ: Prentice Hall).

van Holde, K. E., Johnson, W. C., and Ho, P.S. (2005). *Principles of Physical Biochemistry* 2nd ed. (Upper Saddle River, NJ: Prentice Hall).

Cells

Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2004). *Essential Cell Biology* (New York: Garland).

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Kreiger, M., Scott, M. P., Zipursky, L., and Darnell, J. (2003). *Molecular Cell Biology*, 5th ed. (New York: Scientific American Books).

Goodsell, D. S. (1993). *The Machinery of Life* (New York: Springer-Verlag).

Evolution and the Diversity of Life

Doolittle, W. F. (2000). Uprooting the tree of life. *Sci. Am.* 282(2):90–95.

Doolittle, W. F. (2009). Eradicating topological thinking in prokaryotic systematics and evolution. *Cold Spr. Hbr. Symp. Quant. Biol.*

Margulis, L., and Schwartz, K.V. (1998). *Five Kingdoms*, 3rd ed. (New York: W.H. Freeman).

Graur, D., and Li, W.-H. (2000). *Fundamentals of Molecular Evolution* (Sunderland, MA: Sinauer).

Sapp, J. (Ed.) (2005). *Microbial Phylogeny and Evolution: Concepts and Controversies*. (Oxford, UK: Oxford University Press).

Sapp, J. (2009) *The New Foundations of Evolution*. (Oxford, UK: Oxford University Press).

History of Science

Kohler, R. E. (1975). The History of Biochemistry, a Survey. *J. Hist. Biol* 8:275–318.

Water

Life on Earth is often described as a carbon-based phenomenon but it would be equally correct to refer to it as a water-based phenomenon. Life probably originated in water more than three billion years ago and all living cells still depend on water for their existence. Water is the most abundant molecule in most cells accounting for 60% to 90% of the mass of the cell. The exceptions are cells from which water is expelled such as those in seeds and spores. Seeds and spores can lie dormant for long periods of time until they are revived by the reintroduction of water.

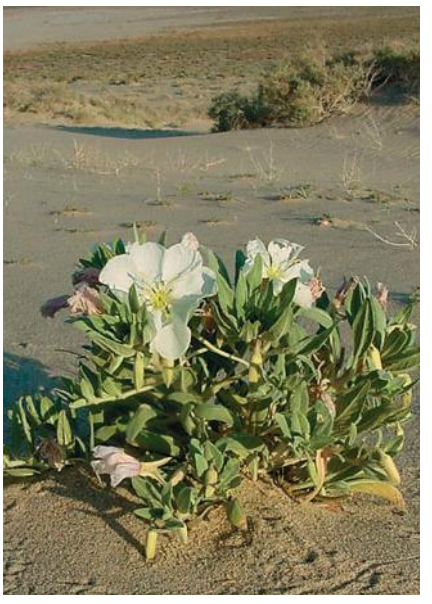
Life spread from the oceans to the continents about 500 million years ago. This major transition in the history of life required special adaptations to enable terrestrial life to survive in an environment where water was less plentiful. You will encounter many of these adaptations in the rest of this book.

An understanding of water and its properties is important to the study of biochemistry. The macromolecular components of cells—proteins, polysaccharides, nucleic acids, and lipids—assume their characteristic shapes in response to water. For example, some types of molecules interact extensively with water and, as a result, are very soluble while other molecules do not dissolve easily in water and tend to associate with each other in order to avoid water. Much of the metabolic machinery of cells has to operate in an aqueous environment because water is an essential solvent.

We begin our detailed study of the chemistry of life by examining the properties of water. The physical properties of water allow it to act as a solvent for ionic and other polar substances, and the chemical properties of water allow it to form weak bonds with other compounds, including other water molecules. The chemical properties of water are also related to the functions of macromolecules, entire cells, and organisms. These interactions are important sources of structural stability in macromolecules and large cellular structures. We will see how water affects the interactions of substances that have low solubility in water. We will examine the ionization of water and discuss acid–base chemistry—topics that are the foundation for understanding the molecules and processes that we will encounter in subsequent chapters. It's important to keep in mind that water is not just an inert solvent; it is also a substrate for many cellular reactions.

There is nothing softer and weaker than water, And yet there is nothing better for attacking hard and strong things. For this reason there is no substitute for it.

—Lao-Tzu (c. 550 BCE)



▲ **Eureka Dunes evening primrose (*Oenothera californica*)** This species only grows in the sand dunes of Death Valley National Park in California. It has evolved special mechanisms for conserving water.

Top: Earth from space. The earth is a watery planet and water plays a central role in the chemistry of all life.

2.1 The Water Molecule Is Polar

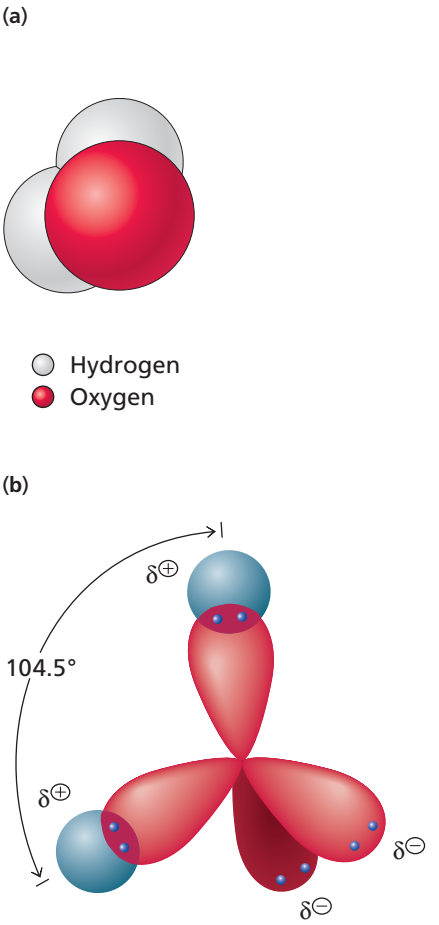
A water molecule (H_2O) is V-shaped (Figure 2.1a) and the angle between the two covalent (O—H) bonds is 104.5°. Some important properties of water arise from its angled shape and the intermolecular bonds that it can form. An oxygen atom has eight electrons and its nucleus has eight protons and eight neutrons. There are two electrons in the inner shell and six electrons in the outer shell. The outer shell can potentially accommodate four pairs of electrons in one *s* orbital and three *p* orbitals. However, the structure of water and its properties can be better explained by assuming that the electrons in the outer shell occupy four sp^3 hybrid orbitals. Think of these four orbitals as occupying the four corners of a tetrahedron that surrounds the central atom of oxygen. Two of the sp^3 hybrid orbitals contain a pair of electrons and the other two each contain a single electron. This means that oxygen can form covalent bonds with other atoms by sharing electrons to fill these single electron orbitals. In water the covalent bonds involve two different hydrogen atoms each of which shares its single electron with the oxygen atom. In Figure 2.1b each electron is indicated by a blue dot showing that each sp^3 hybrid orbital of the oxygen atom is occupied by two electrons including those shared with the hydrogen atoms. The inner shell of the hydrogen atom is also filled because of these two shared electrons in the covalent bond.

The H—O—H bond angle in free water molecules is 104.5° but if the electron orbitals were really pointing to the four corners of a tetrahedron, the angle would be 109.5°. The usual explanation for this difference is that there is strong repulsion between the lone electron pairs and this repulsion pushes the covalent bond orbitals closer together, reducing the angle from 109.5° to 104.5°.

Oxygen atoms are more electronegative than hydrogen atoms because an oxygen nucleus attracts electrons more strongly than the single proton in the hydrogen nucleus. As a result, an uneven distribution of charge occurs within each O—H bond of the water molecule with oxygen bearing a partial negative charge ($\delta^{\ominus}$) and hydrogen bearing a partial positive charge ($\delta^{\oplus}$). This uneven distribution of charge within a bond is known as a dipole and the bond is said to be polar.

The polarity of a molecule depends both on the polarity of its covalent bonds and its geometry. The angled arrangement of the polar O—H bonds of water creates a permanent dipole for the molecule as a whole as shown in Figure 2.2a. A molecule of ammonia also contains a permanent dipole (Figure 2.2b) Thus, even though water and gaseous ammonia are electrically neutral, both molecules are polar. The high solubility of the polar ammonia molecules in water is facilitated by strong interactions with the polar water molecules. The solubility of ammonia in water demonstrates the principle that "like dissolves like."

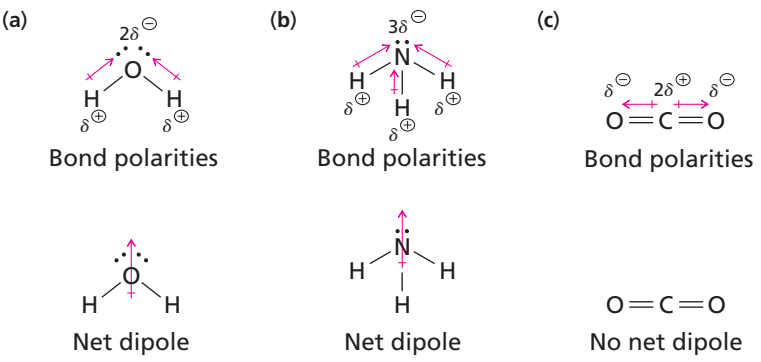
Not all molecules are polar; for example, carbon dioxide also contains polar covalent bonds but the bonds are aligned with each other and oppositely oriented so the polarities cancel each other (Figure 2.2c). As a result, carbon dioxide has no net dipole and is much less soluble in water than ammonia.



▲ **Figure 2.1 A water molecule.** **(a)** Space-filling structure of a water molecule. **(b)** Angle between the covalent bonds of a water molecule. Two of the sp^3 hybrid orbitals of the oxygen atom participate in covalent bonds with *s* orbitals of hydrogen atoms. The other two sp^3 orbitals are occupied by lone pairs of electrons.

KEY CONCEPT

Polar molecules are molecules with an unequal distribution of charge so that one end of the molecules is more negative and another end is more positive.



◄ **Figure 2.2**
Polarity of small molecules. (a) The geometry of the polar covalent bonds of water creates a permanent dipole for the molecule with the oxygen bearing a partial negative charge (symbolized by $2\delta^{\ominus}$) and each hydrogen bearing a partial positive charge (symbolized by $\delta^{\oplus}$). **(b)** The pyramidal shape of a molecule of ammonia also creates a permanent dipole. **(c)** The polarities of the collinear bonds in carbon dioxide cancel each other. Therefore, CO_2 is not polar. (Arrows depicting dipoles point toward the negative charge with a cross at the positive end.)

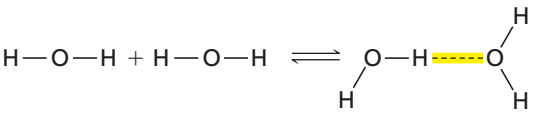
2.2 Hydrogen Bonding in Water

One of the important consequences of the polarity of the water molecule is that water molecules attract one another. The attraction between one of the slightly positive hydrogen atoms of one water molecule and the slightly negative electron pairs in one of the sp^3 hybrid orbitals produces a hydrogen bond (Figure 2.3). In a **hydrogen bond** between two water molecules the hydrogen atom remains covalently bonded to its oxygen atom, the hydrogen donor. At the same time, it is attracted to another oxygen atom, called the hydrogen acceptor. In effect, the hydrogen atom is being shared (unequally) between the two oxygen atoms. The distance from the hydrogen atom to the acceptor oxygen atom is about twice the length of the covalent bond.

KEY CONCEPT

Hydrogen bonds form when a hydrogen atom with a partially positive charge ($\delta^{\oplus}$) is shared between two electronegative atoms ($2\delta^{\ominus}$). Hydrogen bonds are much weaker than covalent bonds.

Water is not the only molecule capable of forming hydrogen bonds; these interactions can occur between any electronegative atom and a hydrogen atom attached to another electronegative atom. (We will examine other examples of hydrogen bonding in Section 2.5B.) Hydrogen bonds are much weaker than typical covalent bonds. The strength of hydrogen bonds in water and in solutions is difficult to measure directly but it is estimated to be about 20 kJ mol^{-1}.



$$\Delta H_f = -20 \text{ kJ mol}^{-1} \qquad (2.1)$$

About 20 kJ mol^{-1} of heat is given off when hydrogen-bonded water molecules form in water under standard conditions. (Recall that standard conditions are 1 atm pressure and a temperature of 25°C.) This value is the standard enthalpy of formation (ΔH_f). It means that the change in enthalpy when hydrogen bonds form is about –20 kJ per mole of water. This is equivalent to saying that +20 kJ mol^{-1} of heat energy is required to disrupt hydrogen bonds between water molecules—the reverse of the reaction shown in Reaction 2.1. This value depends on the type of hydrogen bond. In contrast, the energy required to break a covalent O—H bond in water is about 460 kJ mol^{-1}, and the energy required to break a covalent C—H bond is about 410 kJ mol^{-1}. Thus, the strength of hydrogen bonds is less than 5% of the strength of typical covalent bonds. Hydrogen bonds are weak interactions compared to covalent bonds.

Orientation is important in hydrogen bonding. A hydrogen bond is most stable when the hydrogen atom and the two electronegative atoms associated with it (the two oxygen atoms, in the case of water) are aligned, or nearly in line, as shown in Figure 2.3. Water molecules are unusual because they can form four O—H—O aligned hydrogen bonds with up to four other water molecules (Figure 2.4). They can donate each of their two hydrogen atoms to two other water molecules and accept two hydrogen atoms from two other water molecules. Each hydrogen atom can participate in only one hydrogen bond.

The three-dimensional interactions of liquid water are difficult to study but much has been learned by examining the structure of ice crystals (Figure 2.5). In the common form of ice, every molecule of water participates in four hydrogen bonds, as expected. Each of the hydrogen bonds points to the oxygen atom of an adjacent water molecule and these four adjacent hydrogen-bonded oxygen atoms occupy the vertices of a tetrahedron. This arrangement is consistent with the structure of water shown in Figure 2.1

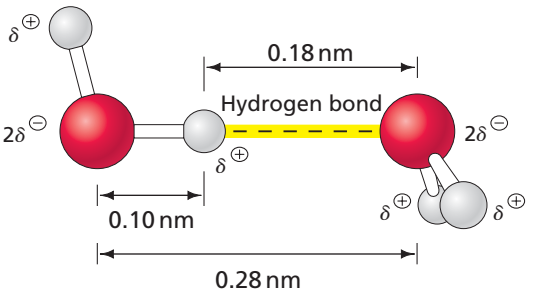


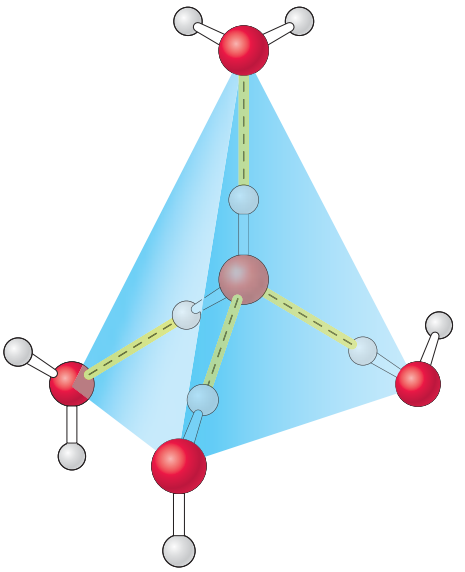
Figure 2.3 ►
Hydrogen bonding between two water molecules. A partially positive ($\delta^{\oplus}$) hydrogen atom of one water molecule attracts the partially negative ($2\delta^{\ominus}$) oxygen atom of a second water molecule, forming a hydrogen bond. The distances between atoms of two water molecules in ice are shown. Hydrogen bonds are indicated by dashed lines highlighted in yellow, as shown here and throughout the book.

except that the bond angles are all equal (109.5°). This is because the polarity of individual water molecules, which distorts the bond angles, is canceled by the presence of hydrogen bonds. The average energy required to break each hydrogen bond in ice has been estimated to be 23 kJ mol^{-1}, making those bonds a bit stronger than those formed in water.

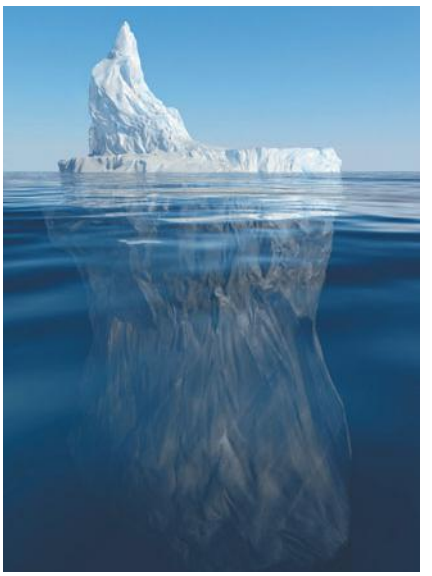
The ability of water molecules in ice to form four hydrogen bonds and the strength of these hydrogen bonds give ice an unusually high melting point because a large amount of energy, in the form of heat, is required to disrupt the hydrogen-bonded lattice of ice. When ice melts most of the hydrogen bonds are retained by liquid water. Each molecule of liquid water can form up to four hydrogen bonds with its neighbors but most participate in only two or three at any given moment. This means that the structure of liquid water is less ordered than that of ice. The fluidity of liquid water is primarily a consequence of the constantly fluctuating pattern of hydrogen bonding as hydrogen bonds break and re-form. At any given time there will be many water molecules participating in two, three, or four hydrogen bonds with other water molecules. There will also be many that participate in only one hydrogen bond or none at all. This is a dynamic structure—the average hydrogen bond lifetime in water is only 10 picoseconds (10^{-11} s).

The density of most substances increases upon freezing as molecular motion slows and tightly packed crystals form. The density of water also increases as it cools—until it reaches a maximum of 1.000 g ml^{-1} at 4°C (277 K). (This value is not a coincidence. Grams are *defined* as the weight of 1 milliliter of water at 4°C.) Water expands as the temperature drops below 4°C. This expansion is caused by the formation of the more open hydrogen-bonded ice crystal in which each water molecule is hydrogen-bonded rigidly to four others. As a result ice is slightly less dense (0.924 g ml^{-1}) than liquid water whose molecules can move enough to pack more closely. Because ice is less dense than liquid water it floats and water freezes from the top down. This has important biological implications since a layer of ice on a pond insulates the creatures below from extreme cold.

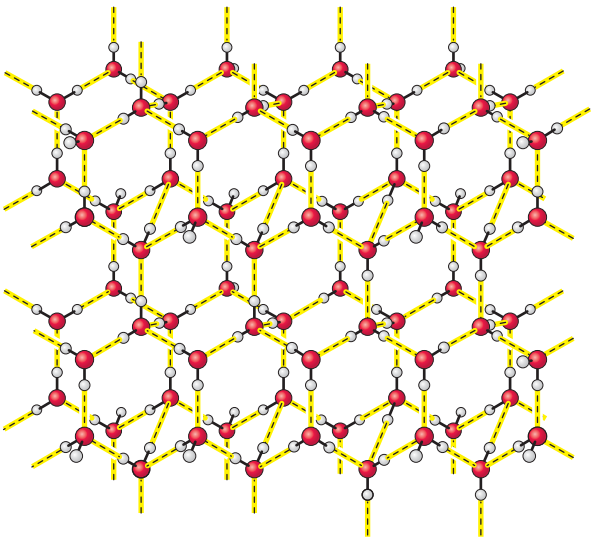
Two additional properties of water are related to its hydrogen-bonding characteristics—its specific heat and its heat of vaporization. The specific heat of a substance is the amount of heat needed to raise the temperature of 1 gram of the substance by 1°C. This property is also called the heat capacity. In the case of water, a relatively large amount of heat is required to raise the temperature because each water molecule participates in multiple hydrogen bonds that must be broken in order for the kinetic energy of the water molecules to increase. The abundance of water in the cells and tissues of all large multicellular organisms means that temperature fluctuations within cells are minimized.



▲ **Figure 2.4**
Hydrogen bonding by a water molecule. A water molecule can form up to four hydrogen bonds: the oxygen atom of a water molecule is the hydrogen acceptor for two hydrogen atoms, and each O—H group serves as a hydrogen donor.



▲ **Icebergs.** Ice floats because it is less dense than water. However, it is only slightly less dense than water so most of the mass of floating ice lies underwater.



◀ **Figure 2.5**
Structure of ice. Water molecules in ice form an open hexagonal lattice in which every water molecule is hydrogen-bonded to four others. The geometrical regularity of these hydrogen bonds contributes to the strength of the ice crystal. The hydrogen-bonding pattern of ice is more regular than that of water. The absolute structure of liquid water has not been determined.

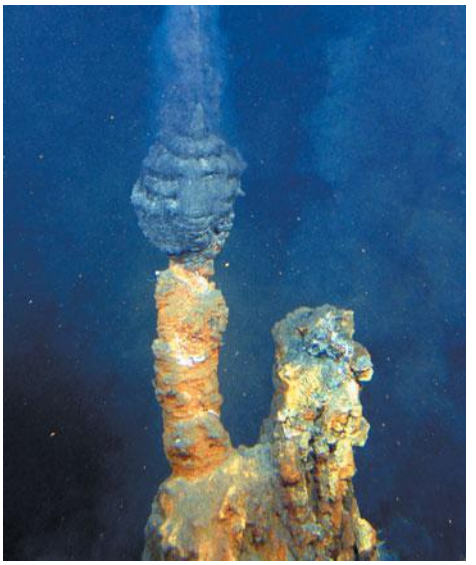
BOX 2.1 EXTREME THERMOPHILES

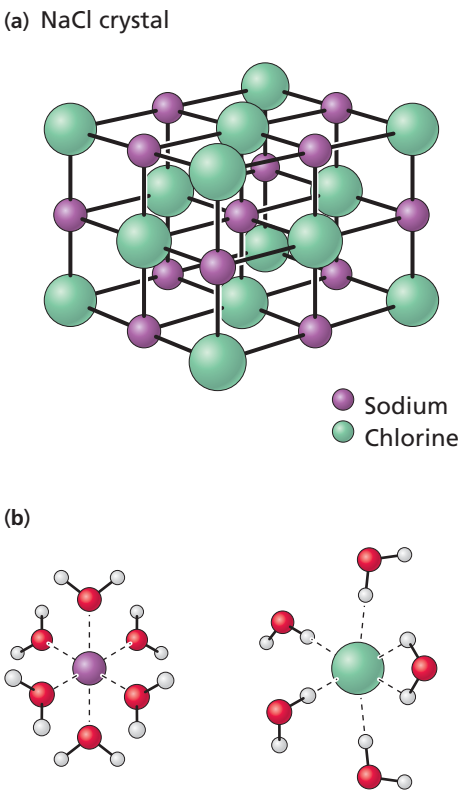
Some species can grow and reproduce at temperatures very close to 0°C, or even lower. There are cold-blooded fish, for example, that survive at ocean temperatures below 0°C (salt lowers the freezing point of water).

At the other extreme are bacteria that live in hot springs where the average temperature is above 80°C. Some bacteria inhabit the environment around deep ocean thermal vents (black smokers) where the average temperature is more than 100°C. (The high pressure at the bottom of the ocean raises the boiling point of water.)

The record for extreme thermophiles is Strain 121, a species of archaebacteria that grows and reproduces at 121°C! These extreme thermophiles are among the earliest branching lineages on the web of life. It's possible that the first living cells arose near deep ocean vents.

Deep ocean hydrothermal vent. ►





▲ Figure 2.6
Dissolution of sodium chloride (NaCl) in water. **(a)** The ions of crystalline sodium chloride are held together by electrostatic forces. **(b)** Water weakens the interactions between the positive and negative ions and the crystal dissolves. Each dissolved $Na^{\oplus}$ and $Cl^{\ominus}$ is surrounded by a solvation sphere. Only one layer of solvent molecules is shown. Interactions between ions and water molecules are indicated by dashed lines.

This feature is of critical biological importance since the rates of most biochemical reactions are sensitive to temperature.

The heat of vaporization of water (~2260 J g^{-1}) is also much higher than that of many other liquids. A large amount of heat is required to convert water from a liquid to a gas because hydrogen bonds must be broken to permit water molecules to dissociate from one another and enter the gas phase. Because the evaporation of water absorbs so much heat, perspiration is an effective mechanism for decreasing body temperature.

2.3 Water Is an Excellent Solvent

The physical properties of water combine to make it an excellent solvent. We have already seen that water molecules are polar and this property has important consequences, as we will see below. In addition, water has a low intrinsic viscosity that does not greatly impede the movement of dissolved molecules. Finally, water molecules themselves are small compared to some other solvents such as ethanol and benzene. The small size of water molecules means that many of them can associate with solute particles to make them more soluble.

A. Ionic and Polar Substances Dissolve in Water

Water can interact with and dissolve other polar compounds and compounds that ionize. Ionization is associated with the gain or loss of an electron, or an H^+ ion, giving rise to an atom or a molecule that carries a net charge. Molecules that can dissociate to form ions are called **electrolytes**. Substances that readily dissolve in water are said to be **hydrophilic**, or water loving. (We will discuss hydrophobic, or water fearing, substances in the next section.)

Why are electrolytes soluble in water? Recall that water molecules are polar. This means they can align themselves around electrolytes so that the negative oxygen atoms of the water molecules are oriented toward the cations (positively charged ions) of the electrolytes and the positive hydrogen atoms are oriented toward the anions (negatively charged ions). Consider what happens when a crystal of sodium chloride (NaCl) dissolves in water (Figure 2.6) The polar water molecules are attracted to the charged ions in the crystal. The attractions result in sodium and chloride ions on the surface of the

crystal dissociating from one another and the crystal begins to dissolve. Because there are many polar water molecules surrounding each dissolved sodium and chloride ion, the interactions between the opposite electric charges of these ions become much weaker than they are in the intact crystal. As a result of its interactions with water molecules, the ions of the crystal continue to dissociate until the solution becomes saturated. At this point, the ions of the dissolved electrolyte are present at high enough concentrations for them to again attach to the solid electrolyte, or crystallize, and an equilibrium is established between dissociation and crystallization.

BOX 2.2 BLOOD PLASMA AND SEAWATER

There was a time when people believed that the ionic composition of blood plasma resembled that of seawater. This was supposed to be evidence that primitive organisms lived in the ocean and land animals evolved a system of retaining the ocean-like composition of salts.

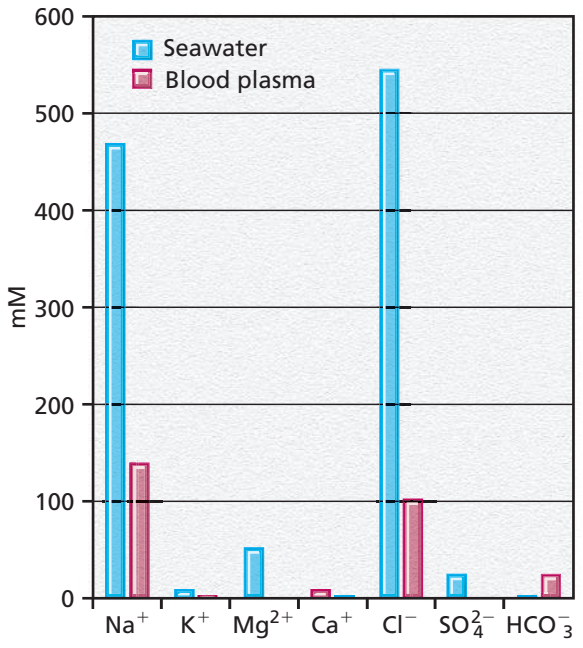
Careful studies of salt concentrations in the early 20th century revealed that the concentration of salts in the ocean were much higher than in blood plasma. Some biochemists tried to explain this discrepancy by postulating that the composition of blood plasma didn't resemble the seawater of today but it did resemble the composition of ancient seawater from several hundred million years ago when multicellular animals arose.

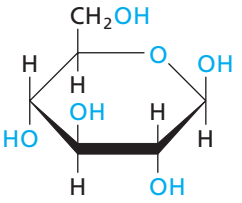
We now know that the saltiness of the ocean hasn't changed very much from the time it first formed over three billion years ago. There is no direct connection between the saltiness of blood plasma and seawater. Not only are the overall concentrations of the major ions (Na^+, K^+, and Cl^-) very different but the relative concentrations of various other ionic species are even more different.

The ionic composition of blood plasma is closely mimicked by Ringer's solution, which also contains lactate as a carbon source. Ringer's solution can be used as a temporary substitute for blood plasma when a patient has suffered blood loss or dehydration.

	Blood plasma	Ringer's
Na^+	140 mM	130 mM
K^+	4 mM	4 mM
Cl^-	103 mM	109 mM
Ca^+	2 mM	2 mM
lactate	5 mM	28 mM

▼ The concentrations of various ions in seawater (blue) and human blood plasma (red) are compared. Seawater is much saltier and contains much higher proportions of magnesium and sulfates. Blood plasma is enriched in bicarbonate (see Section 2.10).





▲ **Figure 2.7**
Structure of glucose. Glucose contains five hydroxyl groups and a ring oxygen, each of which can form hydrogen bonds with water.

Each dissolved $Na^{\oplus}$ attracts the negative ends of several water molecules whereas each dissolved $Cl^{\ominus}$ attracts the positive ends of several water molecules (Figure 2.6b). The shell of water molecules that surrounds each ion is called a solvation sphere and it usually contains several layers of solvent molecules. A molecule or ion surrounded by solvent molecules is said to be **solvated**. When the solvent is water, such molecules or ions are said to be **hydrated**.

Electrolytes are not the only hydrophilic substances that are soluble in water. Any polar molecule will have a tendency to become solvated by water molecules. In addition, the solubility of many organic molecules is enhanced by formation of hydrogen bonds with water molecules. Ionic organic compounds such as carboxylates and protonated amines owe their solubility in water to their polar functional groups. Other groups that confer water solubility include amino, hydroxyl, and carbonyl groups. Molecules containing such groups disperse among water molecules with their polar groups forming hydrogen bonds with water.

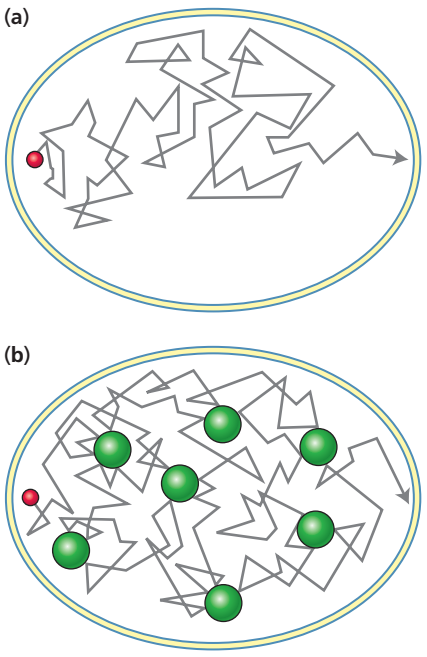
An increase in the number of polar groups in an organic molecule increases its solubility in water. The carbohydrate glucose contains five hydroxyl groups and a ring oxygen (Figure 2.7) and is very soluble in water (up to 83 grams of glucose can dissolve in 100 milliliters of water at 17.5°C). Each oxygen atom of glucose can form hydrogen bonds with water. We will see in other chapters that the attachment of carbohydrates to some otherwise poorly soluble molecules, including lipids and the bases of nucleosides, increases their solubility.

B. Cellular Concentrations and Diffusion

The inside of a cell can be very crowded as suggested by David Goodsell's drawings (Figure 1.17). Consequently, the behavior of solutes in the cytoplasm will be different from their behavior in a simple solution of water. One of the most important differences is reduction of the diffusion rate inside cells.

There are three reasons why solutes diffuse more slowly in cytoplasm.

1. The viscosity of cytoplasm is higher than that of water due to the presence of many solutes such as sugars. This is not an important factor because recent measurements suggest that the viscosity of cytoplasm is only slightly greater than water even in densely packed organelles.
2. Charged molecules bind transiently to each other inside cells and this restricts their mobility. These binding effects have a small but significant effect on diffusion rates.
3. Collisions with other molecules inhibit diffusion due to an effect called **molecular crowding**. This is the main reason why diffusion is slowed in the cytoplasm.



▲ **Figure 2.8**
Diffusion. (a) If the cytoplasm were simply made up of water, a small molecule (red) would diffuse from one end of a cell to the other via a random walk. **(b)** The average time could be about 10 times longer in a crowded cytoplasm, with larger molecules (green).

For small molecules, the diffusion rate inside cells is never more than one-quarter the rate in pure water. For large molecules, such as proteins, the diffusion rate in the cytoplasm may be slowed to about 5% to 10% of the rate in water. This slowdown is due largely to molecular crowding.

For an individual molecule, the rate of diffusion in water at 20°C is described by the diffusion coefficient ($D_{20,w}$). For the protein myoglobin, $D_{20,w} = 11.3 \times 10^{-7}$ cm^2 s^{-1}. From this value we can calculate that the average time to diffuse from one end of a cell to the other (~10 μm) is about 0.44 seconds.

But this diffusion time represents the diffusion time in pure water. In the crowed environment of a typical cell it could take about 10 times longer (4 s). The slower rate is due to the fact that a protein like myoglobin will be constantly bumping into other large molecules. Nevertheless, 4 seconds is still a short time. It means that most molecules, including smaller metabolites and ions, will encounter each other frequently inside a typical cell (Figure 2.8). Recent direct measurements of diffusion inside cells reveal that the effects of molecular crowding are less significant than we used to believe.

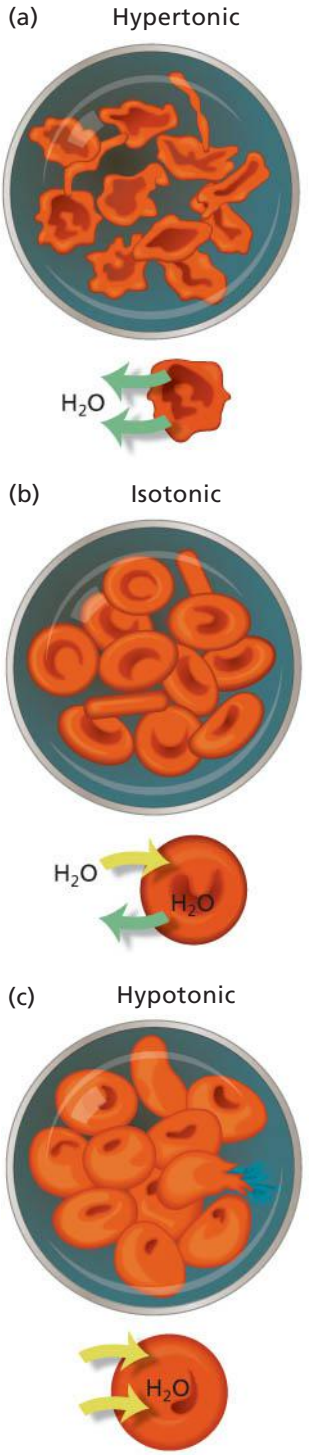
C. Osmotic Pressure

If a solvent-permeable membrane separates two solutions that contain different concentrations of dissolved substances, or solutes, then molecules of *solvent* will diffuse from the less concentrated solution to the more concentrated solution in a process

called **osmosis**. The pressure required to prevent the flow of solvent is called **osmotic pressure**. The osmotic pressure of a solution depends on the total molar concentration of solute, not on its chemical nature.

Water-permeable membranes separate the cytosol from the external medium. The compositions of intracellular solutions are quite different from those of extracellular solutions with some compounds being more concentrated and some less concentrated inside cells. In general, the concentrations of solutes inside the cell are much higher than their concentrations in the aqueous environment outside the cell. Water molecules tend to move across the cell membrane in order to enter the cell and dilute the solution inside the cell. The influx of water causes the cell's volume to increase but this expansion is limited by the cell membrane. In extreme cases, such as when red blood cells are diluted in pure water, the internal pressure causes the cells to burst. Some species (e.g., plants and bacteria) have rigid cell walls that prevent the membrane expansion. These cells can develop high internal pressures.

Most cells use several strategies to keep the osmotic pressure from becoming too great and bursting the cell. One strategy involves condensing many individual molecules into a macromolecule. For example, animal cells that store glucose package it as a polymer called glycogen which contains about 50,000 glucose residues. If the glucose molecules were not condensed into a single glycogen molecule the influx of water necessary to dissolve each glucose molecule would cause the cell to swell and burst. Another strategy is to surround cells with an isotonic solution that negates a net efflux or influx of water. Blood plasma, for example, contains salts and other molecules that mimic the osmolarity inside red blood cells (see Box 2.2).



▲ Hypertonic (a), isotonic (b) and hypotonic (c) red blood cells.

2.4 Nonpolar Substances Are Insoluble in Water

Hydrocarbons and other nonpolar substances have very low solubility in water because water molecules tend to interact with other water molecules rather than with nonpolar molecules. As a result, water molecules exclude nonpolar substances forcing them to associate with each other. For example, tiny oil droplets that are vigorously dispersed in water tend to coalesce to form a single drop thereby minimizing the area of contact between the two substances. This is why the oil in a salad dressing separates if you let it sit for any length of time before putting it on your salad.

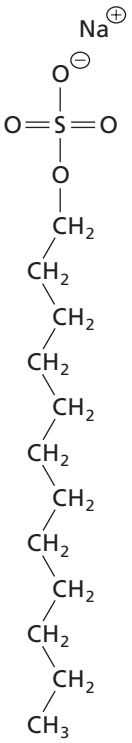
Nonpolar molecules are said to be **hydrophobic**, or water fearing, and this phenomenon of exclusion of nonpolar substances by water is called the **hydrophobic effect**. The hydrophobic effect is critical for the folding of proteins and the self-assembly of biological membranes.

The number of polar groups in a molecule affects its solubility in water. Solubility also depends on the ratio of polar to nonpolar groups in a molecule. For example, one-, two-, and three-carbon alcohols are miscible with water but larger hydrocarbons with single hydroxyl groups are much less soluble in water (Table 2.1). In the larger

Table 2.1 Solubilities of short-chain alcohols in water

Alcohol	Structure	Solubility in water (mol/100 g H_2O at 20°C)[a]
Methanol	CH_3OH	∞
Ethanol	CH_3CH_2OH	∞
Propanol	$CH_3(CH_2)_2OH$	∞
Butanol	$CH_3(CH_2)_3OH$	0.11
Pentanol	$CH_3(CH_2)_4OH$	0.030
Hexanol	$CH_3(CH_2)_5OH$	0.0058
Heptanol	$CH_3(CH_2)_6OH$	0.0008

[a] Infinity (∞) indicates that there is no limit to the solubility of the alcohol in water.



▲ Figure 2.9
Sodium dodecyl sulfate (SDS), a synthetic detergent.

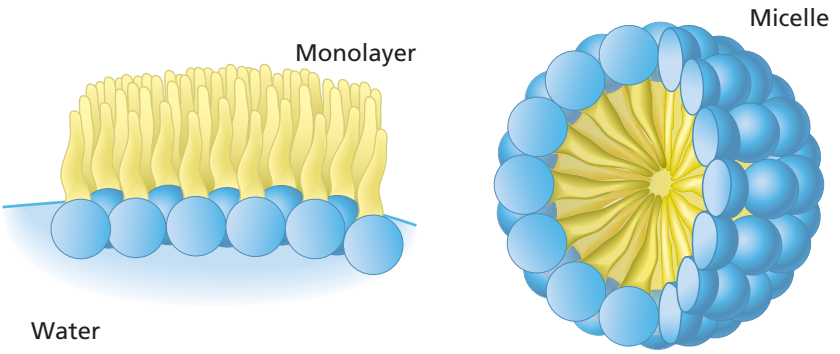
molecules, the properties of the nonpolar hydrocarbon portion of the molecule override those of the polar alcohol group and limit solubility.

Detergents, sometimes called surfactants, are molecules that are both hydrophilic and hydrophobic. They usually have a hydrophobic chain at least 12 carbon atoms long and an ionic or polar end. Such molecules are said to be **amphipathic**. Soaps, which are alkali metal salts of long-chain fatty acids are one type of detergent. The soap sodium palmitate ($CH_3(CH_2)_{14}COO^{\ominus}Na^{\oplus}$), for example, contains a hydrophilic carboxylate group and a hydrophobic tail. One of the synthetic detergents most commonly used in biochemistry is sodium dodecyl sulfate (SDS) which contains a 12-carbon tail and a polar sulfate group (Figure 2.9).

The hydrocarbon portion of a detergent is soluble in nonpolar organic substances and its polar group is soluble in water. When a detergent is spread on the surface of water a monolayer forms in which the hydrophobic, nonpolar tails of the detergent molecules extend into the air groups of detergent molecules aggregate into micelles while the hydrophilic, ionic heads are hydrated, extending into the water (Figure 2.10). When a sufficiently high concentration of detergent is dispersed in water rather than layered on the surface. In one common form of micelle, the nonpolar tails of the detergent molecules associate with one another in the center of the structure minimizing contact with water molecules. Because the tails are flexible, the core of a micelle is liquid hydrocarbon. The ionic heads project into the aqueous solution and are therefore hydrated. Small, compact micelles may contain about 80 to 100 detergent molecules.

The cleansing action of soaps and other detergents derives from their ability to trap water-insoluble grease and oils within the hydrophobic interiors of micelles. SDS and similar synthetic detergents are common active ingredients in laundry detergents. The suspension of nonpolar compounds in water by their incorporation into micelles is termed *solubilization*. Solubilizing nonpolar molecules is a different process than dissolving a polar compound. A number of the structures that we will encounter later in this book, including proteins and biological membranes, resemble micelles in having hydrophobic interiors and hydrophilic surfaces.

Some dissolved ions such as $SCN^{\ominus}$ (thiocyanate) and $ClO_4^{\ominus}$ (perchlorate) are called **chaotropes**. These ions are poorly solvated compared to ions such as $NH4^{\oplus}$, $SO_4^{2\ominus}$, and $H_2PO_4^{\ominus}$. Chaotropes enhance the solubility of nonpolar compounds in water by disordering the water molecules (there is no general agreement on how chaotropes do this). We will encounter other examples of chaotropic agents such as the guanidinium ion and the nonionic compound urea when we discuss denaturation and the three-dimensional structures of proteins and nucleic acids.



▲ Figure 2.10
Cross-sectional views of structures formed by detergents in water. Detergents can form monolayers at the air–water interface. They can also form micelles, aggregates of detergent molecules in which the hydrocarbon tails (yellow) associate in the water-free interior and the polar head groups (blue) are hydrated.

2.5 Noncovalent Interactions

So far in this chapter we have introduced two types of noncovalent interactions—hydrogen bonds and hydrophobic interactions. Weak interactions such as these play extremely important roles in determining the structures and functions of macromolecules. Weak forces are also involved in the recognition of one macromolecule by another and in the binding of reactants to enzymes.

There are actually four major noncovalent bonds or forces. In addition to hydrogen bonds and hydrophobicity there are also charge–charge interactions and van der Waals forces. Charge–charge interactions, hydrogen bonds, and van der Waals forces are variations of a more general type of force called **electrostatic interactions**.

A. Charge–Charge Interactions

Charge–charge interactions are electrostatic interactions between two charged particles. These interactions are potentially the strongest noncovalent forces and can extend over greater distances than other noncovalent interactions. The stabilization of NaCl crystals by interionic attraction between the sodium ($Na^{\oplus}$) and chloride ($Cl^{\ominus}$) ions is an example of a charge–charge interaction. The strength of such interactions in solution depends on the nature of the solvent. Since water greatly weakens these interactions, the stability of macromolecules in an aqueous environment is not strongly dependent on charge–charge interactions but they do occur. An example of charge-charge interactions in proteins is when oppositely charged functional groups attract one another. The interaction is sometimes called a **salt bridge** and it's usually buried deep within the hydrophobic interior of a protein where it can't be disrupted by water molecules. The most accurate term for such interactions is **ion pairing**.

Charge–charge interactions are also responsible for the mutual repulsion of similarly charged ionic groups. Charge repulsion can influence the structures of individual biomolecules as well as their interactions with other, like-charged molecules.

In addition to their relatively minor contribution to the stabilization of large molecules, charge–charge interactions play a role in the recognition of one molecule by another. For example, most enzymes have either anionic or cationic sites that bind oppositely charged reactants.

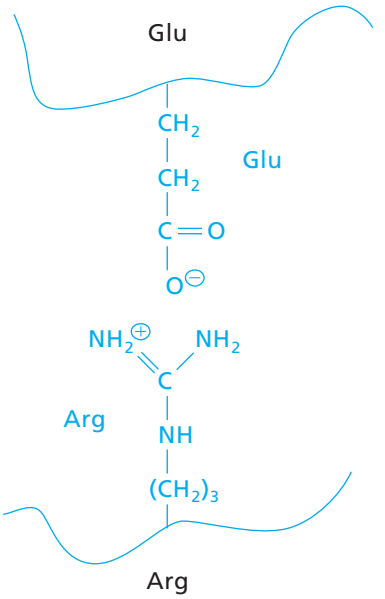
B. Hydrogen Bonds

Hydrogen bonds, which are also a type of electrostatic interaction, occur in many macromolecules and are among the strongest noncovalent forces in biological systems. The strengths of hydrogen bonds such as those between substrates and enzymes and those between the bases of DNA are estimated to be about 25–30 kJ mol^{-1}. These hydrogen bonds are a bit stronger than those formed between water molecules (Section 2.2). Hydrogen bonds in biochemical molecules are strong enough to confer structural stability but weak enough to be broken readily.

In general, when a hydrogen atom is covalently bonded to a strongly electronegative atom, such as nitrogen, oxygen, or sulfur, a hydrogen bond can only form when the hydrogen atom lies approximately 0.2 nm from another strongly electronegative atom with an unshared electron pair. As previously described in the case of hydrogen bonds between water molecules the covalently bonded atom (designated D in Figure 2.11a) is the hydrogen donor and the atom that attracts the proton (designated A in Figure 2.11a) is the hydrogen acceptor. The total distance between the two electronegative atoms participating in a hydrogen bond is typically between 0.27 nm and 0.30 nm. Some common examples of hydrogen bonds are shown in Figure 2.11b.

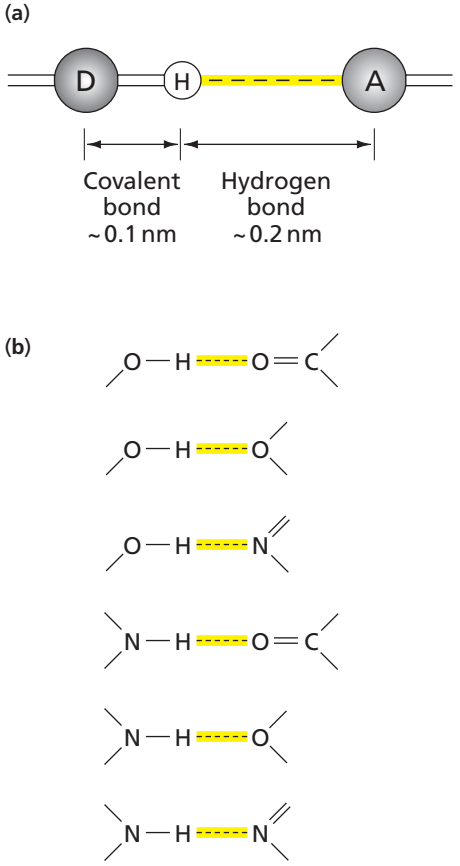
A hydrogen bond has many of the characteristics of a covalent bond but it is much weaker. You can think of a hydrogen bond as a partial sharing of electrons. (Recall that in a true covalent bond a pair of electrons is shared between two atoms.) The three atoms involved in a hydrogen bond are usually aligned to form a straight line where the center of the hydrogen atoms falls directly on a line drawn between the two electronegative

(a)



(b)

▲ **Salt bridges. (a)** One kind of salt bridge. **(b)** Another kind of salt bridge.



▲ Figure 2.11
Hydrogen bonds. (a) Hydrogen bonding between a —D—H group (the hydrogen donor) and an electronegative atom A—(the hydrogen acceptor). A typical hydrogen bond is approximately 0.2 nm long, roughly twice the length of the covalent bond between hydrogen and nitrogen, oxygen, or sulfur. The total distance between the two electronegative atoms participating in a hydrogen bond is therefore approximately 0.3 nm. **(b)** Examples of biologically important hydrogen bonds.

Hydrogen bonding between base pairs in double-stranded DNA makes only a small contribution to the stability of DNA, as described in Section 19.2C.

KEY CONCEPT

Hydrogen bonds between and within biological molecules are easily disrupted by competition with water molecules.

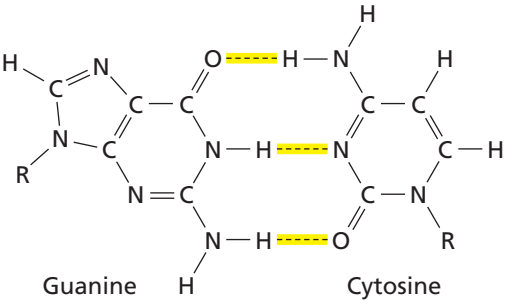


Figure 2.12 ▲
Hydrogen bonding between the complementary bases guanine and cytosine in DNA.

atoms. Small deviations from this alignment are permitted but such hydrogen bonds are weaker than the standard form.

All of the functional groups shown in Figure 2.11 are also capable of forming hydrogen bonds with water molecules. In fact, when they are exposed to water they are far more likely to interact with water molecules because the concentration of water is so high. In order for hydrogen bonds to form between, or within, biochemical macromolecules the donor and acceptor groups have to be shielded from water. In most cases, this shielding occurs because the groups are buried in the hydrophobic interior of the macromolecule where water can't penetrate. In DNA, for example, the hydrogen bonds between complementary base pairs are in the middle of the double helix (Figure 2.12).

C. Van der Waals Forces

The third weak force involves the interactions between permanent or transient dipoles of two molecules. These forces are of short range and small magnitude, about 13 kJ mol^{-1} and 0.8 kJ mol^{-1}, respectively.

These electrostatic interactions are called **van der Waals forces** named after the Dutch physicist Johannes Diderik van der Waals. They only occur when atoms are very close together. Van der Waals forces involve both attraction and repulsion. The attractive forces, also known as London dispersion forces, originate from the infinitesimal dipole generated in atoms by the random movement of the negatively charged electrons around the positively charged nucleus. Thus, van der Waals forces are dipolar, or electrostatic, attractions between the nuclei of atoms or molecules and the electrons of other atoms or molecules. The strength of the interaction between the transiently induced dipoles of nonpolar molecules such as methane is about 0.4 kJ mol^{-1} at an internuclear separation of 0.3 nm. Although they operate over similar distances, van der Waals forces are much weaker than hydrogen bonds.

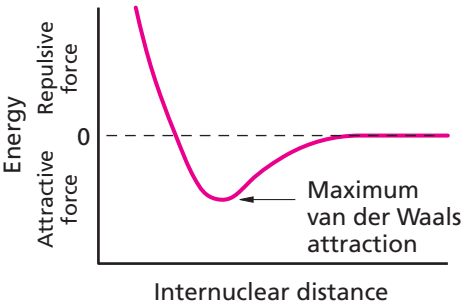
There is also a repulsive component to van der Waals forces. When two atoms are squeezed together the electrons in their orbitals repel each other. The repulsion increases exponentially as the atoms are pressed together and at very close distances it becomes prohibitive.

The sum of the attractive and repulsive components of van der Waals forces yields an energy profile like that in Figure 2.13. At large intermolecular distances the two atoms do not interact and there are no attractive or repulsive forces between them. As the atoms approach each other (moving toward the left in the diagram) the attractive force increases. This attractive force is due to the delocalization of the electron cloud around the atoms. You can picture this as a shift in electrons around one of the atoms such that the electrons tend to localize on the side opposite that of the other approaching atom. This shift creates a local dipole where one side of the atom has a slight positive charge and the other side has a slight negative charge. The side with the small positive charge attracts the other negatively charged atom. As the atoms move even closer together the effect of this dipole diminishes and the overall influence of the negatively charged electron cloud becomes more important. At short distances the atoms repel each other.

The optimal packing distance is the point at which the attractive forces are maximized. This distance corresponds to the energy trough in Figure 2.13 and it is equal to the sum of the van der Waals radii of the two atoms. When the atoms are separated by the sum of their two van der Waals radii they are said to be in van der Waals contact. Typical van der Waals radii of several atoms are shown in Table 2.2.

In some cases, the shift in electrons is influenced by the approach of another atom. This is an induced dipole. In other cases, the delocalization of electrons is a permanent feature of the molecule as we saw in the case of water (Section 2.1). These permanent dipoles also give rise to van der Waals forces.

Although individual van der Waals forces are weak, the clustering of atoms within a protein, nucleic acid, or biological membrane permits formation of a large number of these weak interactions. Once formed, these cumulative weak forces play important roles in maintaining the structures of the molecules. For example, the heterocyclic bases of nucleic acids are stacked one above another in double-stranded DNA. This arrangement is stabilized by a variety of noncovalent interactions, especially van der Waals forces. These forces are collectively known as stacking interactions (see Chapter 19).



▲ **Figure 2.13**
Effect of internuclear separation on van der Waals forces. Van der Waals forces are strongly repulsive at short internuclear distances and very weak at long internuclear distances. When two atoms are separated by the sum of their van der Waals radii, the van der Waals attraction is maximal.

D. Hydrophobic Interactions

The association of a relatively nonpolar molecule or group with other nonpolar molecules is termed a **hydrophobic interaction**. Although hydrophobic interactions are sometimes called hydrophobic "*bonds*," this description is incorrect. Nonpolar molecules don't aggregate because of mutual attraction but because the polar water molecules surrounding them tend to associate with each other rather than with the nonpolar molecules (Section 2.4). For example, micelles (Figure 2.10) are stabilized by hydrophobic interactions.

The hydrogen-bonding pattern of water is disrupted by the presence of a nonpolar molecule. Thus, water molecules surrounding a less polar molecule in solution are more restricted in their interactions with other water molecules. These restricted water molecules are relatively immobile, or ordered, in the same way that molecules at the surface of water are ordered in the familiar phenomenon of surface tension. However, water molecules in the bulk solvent phase are much more mobile, or disordered. In thermodynamic terms, there is a net gain in the combined entropy of the solvent and the nonpolar solute when the nonpolar groups aggregate and water is freed from its ordered state surrounding the nonpolar groups.

Hydrophobic interactions, like hydrogen bonds, are much weaker than covalent bonds but stronger than van der Waals interactions. For example, the energy required to transfer a —CH_2— group from a hydrophobic to an aqueous environment is about 3 kJ mol^{-1}.

Although individual hydrophobic interactions are weak, the cumulative effect of many hydrophobic interactions can have a significant effect on the stability of a macromolecule. The three-dimensional structure of most proteins, for example, is largely determined by hydrophobic interactions formed during the spontaneous folding of the polypeptide chain. Water molecules are bound to the outside surface of the protein but can't penetrate the interior where most of the nonpolar groups are located.

All four of the interactions covered here are individually weak compared to covalent bonds but the combined effect of many such weak interactions can be quite strong. The most important noncovalent interactions in biomolecules are shown in Figure 2.14.

Table 2.2 Van der Waals radii of several atoms

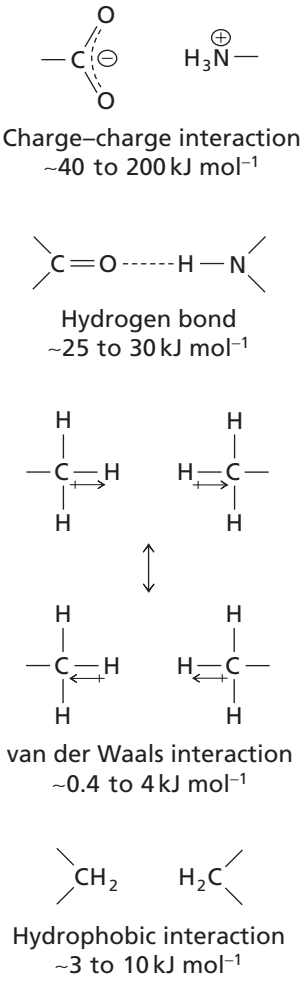
Atom	Radius (nm)
Hydrogen	0.12
Oxygen	0.14
Nitrogen	0.15
Carbon	0.17
Sulfur	0.18
Phosphorus	0.19

KEY CONCEPT

Weak interactions are individually weak but the combined effect of a large number of weak interactions is a significant organizing force.

2.6 Water Is Nucleophilic

In addition to its physical properties, the chemical properties of water are also important in biochemistry because water molecules can react with biological molecules. The electron-rich oxygen atom determines much of water's reactivity in chemical reactions. Electron-rich chemicals are called **nucleophiles** (nucleus lovers) because they seek positively charged (electron-deficient) species called **electrophiles** (electron lovers). Nucleophiles are either negatively charged or have unshared pairs of electrons. They attack



▲ **Figure 2.14**
Typical noncovalent interactions in biomolecules. Charge–charge interactions, hydrogen bonds, and van der Waals interactions are electrostatic interactions. Hydrophobic interactions depend on the increased entropy of the surrounding water molecules rather than on direct attraction between nonpolar groups. For comparison, the dissociation energy for a covalent bond such as C—H or C—C is approximately 340–450 kJ mol^{-1}.

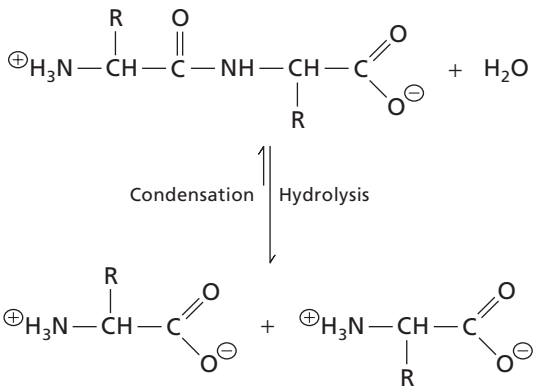


Figure 2.15 ▲
Hydrolysis of a peptide. In the presence of water the peptide bonds in proteins and peptides are hydrolyzed. Condensation, the reverse of hydrolysis, is not thermodynamically favored.

electrophiles during substitution or addition reactions. The most common nucleophilic atoms in biology are oxygen, nitrogen, sulfur, and carbon.

The oxygen atom of water has two unshared pairs of electrons making it nucleophilic. Water is a relatively weak nucleophile but its cellular concentration is so high that one might reasonably expect it to be very reactive. Many macromolecules should be easily degraded by nucleophilic attack by water. This is, in fact, a correct expectation. Proteins, for example, are hydrolyzed, or degraded, by water to release their monomeric units, amino acids (Figure 2.15). The equilibrium for complete hydrolysis of a protein lies far in the direction of degradation; in other words, the ultimate fate of all proteins is destruction by hydrolysis!

If there is so much water in cells then why aren't all biopolymers rapidly degraded? Similarly, if the equilibrium lies toward breakdown, how does biosynthesis occur in an aqueous environment? Cells avoid these problems in several ways. For example, the linkages between the monomeric units of macromolecules, such as the peptide bonds in proteins and the ester linkages in DNA, are relatively stable in solution at cellular pH and temperature in spite of the presence of water. In this case, the stability of linkages refers to their rate of hydrolysis in water and not their thermodynamic stability.

The chemical properties of water combined with its high concentration mean that the Gibbs free energy change for hydrolysis (ΔG) is negative. This means that all hydrolysis reactions are thermodynamically favorable. However, the rate of the reactions inside the cell is so slow that macromolecules are not appreciably degraded by spontaneous hydrolysis during the average lifetime of a cell. It is important to keep in mind the distinction between the preferred direction of a reaction, as indicated by the Gibbs free energy change, and the rate of the reaction, as indicated by the rate constant (Section 1.4D). The key concept is that because of the activation energy there is no direct correlation between the rate of a reaction and the final equilibrium values of the reactants and products.

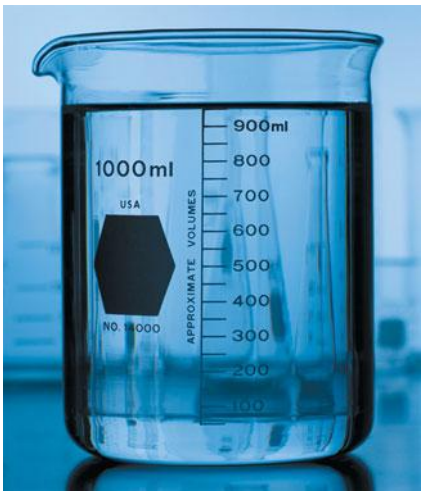
Cells can synthesize macromolecules in an aqueous environment even though condensation reactions—the reverse of hydrolysis—are thermodynamically unfavorable. They do this by using the chemical potential energy of ATP to overcome an unfavorable thermodynamic barrier. Furthermore, the enzymes that catalyze such reactions exclude water from the active site where the synthesis reactions occur. These reactions usually follow two-step chemical pathways that differ from the reversal of hydrolysis. For example, the simple condensation pathway shown in Figure 2.15 is not the pathway that is used in living cells because the presence of high concentrations of water makes the direct condensation reaction extremely unfavorable. In the first synthetic step, which is thermodynamically uphill, the molecule to be transferred reacts with ATP to form a reactive intermediate. In the second step, the activated group is readily

BOX 2.3 THE CONCENTRATION OF WATER

The density of water varies with temperature. It is defined as 1.00000 g/ml at 3.98°C. The density is 0.99987 at 0°C and 0.99707 at 25°C.

The molecular mass of the most common form of water is M_r =18.01056. The concentration of pure water at 3.98°C is 55.5 M (1000 ÷ 18.01).

Many biochemical reactions involve water as either a reactant or a product and the high concentration of water will affect the equilibrium of the reaction.



KEY CONCEPT

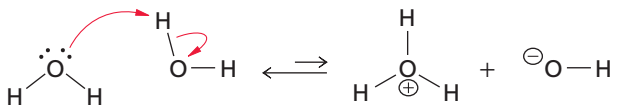
There is a difference between the rate of a reaction and whether it is thermodynamically favorable. Biological molecules are stable because the rate of spontaneous hydrolysis is slow.

transferred to the attacking nucleophile. In Chapter 22 we will see that the reactive intermediate in protein synthesis is an aminoacyl-tRNA that is formed in a reaction involving ATP. The net result of the biosynthesis reaction is to couple the condensation to the hydrolysis of ATP.

The role of ATP in coupled reactions is described in Section 10.7.

2.7 Ionization of Water

One of the important properties of water is its slight tendency to ionize. Pure water contains a low concentration of hydronium ions ($H_3O^{\oplus}$) and an equal concentration of hydroxide ions ($OH^{\ominus}$). The hydronium and hydroxide ions are formed by a nucleophilic attack of oxygen on one of the protons in an adjacent water molecule.



$$H_2O + H_2O \rightleftharpoons H_3O^{\oplus} + OH^{\ominus} \qquad (2.2)$$

The red arrows in Reaction 2.2 show the movement of pairs of electrons. These arrows are used to depict reaction mechanisms and we will encounter many such diagrams throughout this book. One of the free pairs of electrons on the oxygen will contribute to formation of a new O—H covalent bond between the oxygen atom of the hydronium ion and a proton ($H^{\oplus}$) abstracted from a water molecule. An O—H covalent bond is broken in this reaction and the electron pair from that bond remains associated with the oxygen atom of the hydroxide ion.

Note that the atoms in the hydronium ion contain eleven positively charged protons (eight in the oxygen atom and three hydrogen protons) and ten negatively charged electrons (a pair of electrons in the inner orbital of the oxygen atom, one free electron pair associated with the oxygen atom, and three pairs in the covalent bonds). This results in a net positive charge which is why we refer to it as an ion (cation). The positive charge is usually depicted as if it were associated with the oxygen atom but, in fact, it is distributed partially over the hydrogen atoms as well. Similarly, the hydroxide ion (anion) bears a net negative charge because it contains ten electrons whereas the nuclei of the oxygen and hydrogen atoms have a total of only nine positively charged protons.

The ionization reaction is a typical reversible reaction. The protonation and deprotonation reactions take place very quickly. Hydroxide ions have a short lifetime in water and so do hydronium ions. Even water molecules themselves have only a transient existence. The average water molecule is thought to exist for about one millisecond (10^{-3}s) before losing a proton to become a hydroxide ion or gaining a proton to become a hydronium ion. Note that the lifetime of a water molecule is still eight orders of magnitude (10^{8}) greater than the lifetime of a hydrogen bond.

Hydronium ($H_3O^{\oplus}$) ions are capable of donating a proton to another ion. Such proton donors are referred to as **acids** according to the Brønsted–Lowry concept of acids and bases. In order to simplify chemical equations we often represent the hydronium ion as simply $H^{\oplus}$ (free proton or hydrogen ion) to reflect the fact that it is a major source of protons in biochemical reactions. The ionization of water can then be depicted as a simple dissociation of a proton from a single water molecule.

$$H_2O \rightleftharpoons H^{\oplus} + OH^{\ominus} \tag{2.3}$$

Reaction 2.3 is a convenient way to show the ionization of water but it does not reflect the true structure of the proton donor which is actually the hydronium ion. Reaction 2.3 also obscures the fact that the ionization of water is actually a bimolecular reaction involving two separate water molecules as shown in Reaction 2.2. Fortunately, the dissociation of water is a reasonable approximation that does not affect our calculations or our understanding of the properties of water. We will make use of this assumption in the rest of the book.

Hydroxide ions can accept a proton and be converted back into water molecules. Proton acceptors are called **bases**. Water can function as either an acid or a base as Reaction 2.2 demonstrates.

The ionization of water can be analyzed quantitatively. Recall that the concentrations of reactants and products in a reaction will eventually reach an equilibrium where there is no net change in concentration. The ratio of these equilibrium concentrations defines the equilibrium constant (K_{eq}). In the case of ionization of water,

$$K_{eq} = \frac{[H^{\oplus}][OH^{\ominus}]}{H_2O} \qquad K_{eq}[H_2O] = [H^{\oplus}][OH^{\ominus}] \tag{2.4}$$

The density of water varies with the temperature (Box 2.2) and so does the ion product. The differences aren't significant in the temperature ranges that we normally encounter in living cells, so we assume that the value 10^{-14} applies at all temperatures. (See Problem 17 at the end of this chapter.)

The equilibrium constant for the ionization of water has been determined under standard conditions of pressure (1 atm) and temperature (25°C). Its value is 1.8×10^{-16} M. We are interested in knowing the concentrations of protons and hydroxide ions in a solution of pure water since these ions participate in many biochemical reactions. These values can be calculated from Equation 2.4 if we know the concentration of water ($[H_2O]$) at equilibrium. Pure water at 25°C has a concentration of approximately 55.5 M (see Box 2.2). A very small percentage of water molecules will dissociate to form $H^{\oplus}$ and $OH^{\ominus}$ when the ionization reaction reaches equilibrium. This will have a very small effect on the final concentration of water molecules at equilibrium. We can simplify our calculations by assuming that the concentration of water in Equation 2.4 is 55.5 M. Substituting this value, and that of the equilibrium constant, gives

$$(1.8 \times 10^{-16}\ \text{M})(55.5\ \text{M}) = 1.0 \times 10^{-14}\ \text{M}^2 = [H^{\oplus}][OH^{\ominus}] \tag{2.5}$$

The product obtained by multiplying the proton and hydroxide ion concentrations ($[H^{\oplus}][OH^{\ominus}]$) is called the **ion product for water**. This is a constant designated K_w (the ion product constant for water). At 25°C the value of K_w is

$$K_w = [H^{\oplus}][OH^{\ominus}] = 1.0 \times 10^{-14}\ \text{M}^2 \tag{2.6}$$

It is a fortunate coincidence that this is a nice round number rather than some awkward fraction because it makes calculations of ion concentrations much easier. Pure water is

electrically neutral, so its ionization produces an equal number of protons and hydroxide ions $[H^{\oplus}] = [OH]$. In the case of pure water, Equation 2.6 can therefore be rewritten as

$$K_w = [H^{\oplus}]^2 = 1.0 \times 10^{-14}\ M^2 \quad (2.7)$$

Taking the square root of the terms in Equation 2.7 gives

$$[H^{\oplus}] = 1.0 \times 10^{-7}\ M \quad (2.8)$$

Since $[H^{\oplus}] = [OH^{\ominus}]$, the ionization of pure water produces 10^{-7} M $H^{\oplus}$ and 10^{-7} M $OH^{\ominus}$. Pure water and aqueous solutions that contain equal concentrations of $H^{\oplus}$ and $OH^{\ominus}$ are said to be *neutral*. Of course, not all aqueous solutions have equal concentrations of $H^{\oplus}$ and $OH^{\ominus}$. When an acid is dissolved in water $[H^{\oplus}]$ increases and the solution is described as acidic. Note that when an acid is dissolved in water the concentration of protons increases while the concentration of hydroxide ions decreases. This is because the ion product constant for water (K_w) is unchanged (i.e., constant) and the product of the concentrations of $H^{\oplus}$ and $OH^{\ominus}$ must always be 1.0×10^{-14} M^2 under standard conditions (Equation 2.5). Dissolving a base in water decreases $[H^{\oplus}]$ and increases $[OH^{\ominus}]$ above 1.0×10^{-7} M producing a basic, or alkaline, solution.

2.8 The pH Scale

Many biochemical processes—including the transport of oxygen in the blood, the catalysis of reactions by enzymes, and the generation of metabolic energy during respiration or photosynthesis—are strongly affected by the concentration of protons. Although the concentration of $H^{\oplus}$ (or $H_3O^{\oplus}$) in cells is small relative to the concentration of water, the range of $[H^{\oplus}]$ in aqueous solutions is enormous so it is convenient to use a logarithmic quantity called pH as a measure of the concentration of $H^{\oplus}$. **pH** is defined as the negative logarithm of the concentration of $H^{\oplus}$.

$$pH = -\log[H^{\oplus}] = \log\frac{1}{[H^{\oplus}]} \quad (2.9)$$

In pure water $[H^{\oplus}] = [OH^{\ominus}] = 1.0 \times 10^{-7}$ M (Equations 2.7 and 2.8). As mentioned earlier, pure water is said to be "neutral" with respect to total ionic charge since the concentrations of the positively charged hydrogen ions and the negatively charged hydroxide ions are equal. Neutral solutions have a pH value of 7.0 (the negative value of $\log 10^{-7}$ is 7.0). Acidic solutions have an excess of $H^{\oplus}$ due to the presence of dissolved solute that supplies $H^{\oplus}$ ions. In a solution of 0.01 M HCl, for example, the concentration of $H^{\oplus}$ is 0.01 M (10^{-2} M) because HCl dissociates completely to $H^{\oplus}$ and $Cl^{\ominus}$. The pH of such a solution is $-\log 10^{-2} = 2.0$. Thus, the higher the concentration of $H^{\oplus}$, the lower the pH of the solution. The pH scale is logarithmic, so a change in pH of one unit corresponds to a 10-fold change in the concentration of $H^{\oplus}$.

Aqueous solutions can also contain fewer $H^{\oplus}$ ions than pure water resulting in a pH above 7. In a solution of 0.01 M NaOH, for example, the concentration of $OH^{\ominus}$ is 0.01 M (10^{-2} M) because NaOH, like HCl, is 100% dissociated in water. The $H^{\oplus}$ ions derived from the ionization of water will combine with the hydroxide ions from NaOH to re-form water molecules. This affects the equilibrium for the ionization of water (Reaction 2.3). The resulting solution is very basic because of the low concentration of protons. The actual pH can be determined from the ion product of water, K_w (Equation 2.6), by substituting the concentration of hydroxide ions. Since the product of the $OH^{\ominus}$ and $H^{\oplus}$ concentrations is 10^{-14} M it follows that the $H^{\oplus}$ concentration in a solution of 10^{-2} M $OH^{\ominus}$ is 10^{-12} M. The pH of the solution is 12. Table 2.3 shows this relationship between pH and the concentrations of $H^{\oplus}$ and $OH^{\ominus}$.

Basic solutions have pH values greater than 7.0 and acidic solutions have lower pH values. Figure 2.16 illustrates the pH values of various common solutions.

Table 2.3 Relation of $[H^{\oplus}]$ and $[OH^{\ominus}]$ to pH

pH	$[H^{\oplus}]$ (M)	$[OH^{\ominus}]$ (M)
0	1	10^{-14}
1	10^{-1}	10^{-13}
2	10^{-2}	10^{-12}
3	10^{-3}	10^{-11}
4	10^{-4}	10^{-10}
5	10^{-5}	10^{-9}
6	10^{-6}	10^{-8}
7	10^{-7}	10^{-7}
8	10^{-8}	10^{-6}
9	10^{-9}	10^{-5}
10	10^{-10}	10^{-4}
11	10^{-11}	10^{-3}
12	10^{-12}	10^{-2}
13	10^{-13}	10^{-1}
14	10^{-14}	1

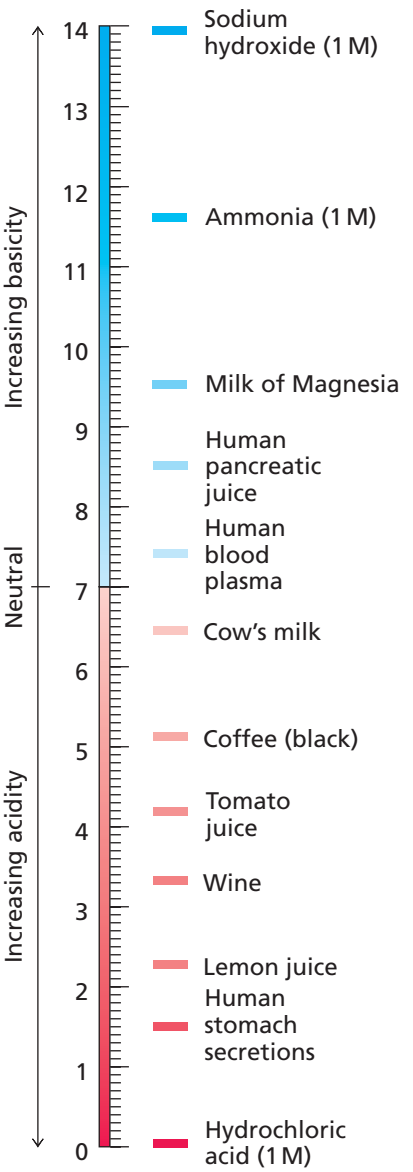


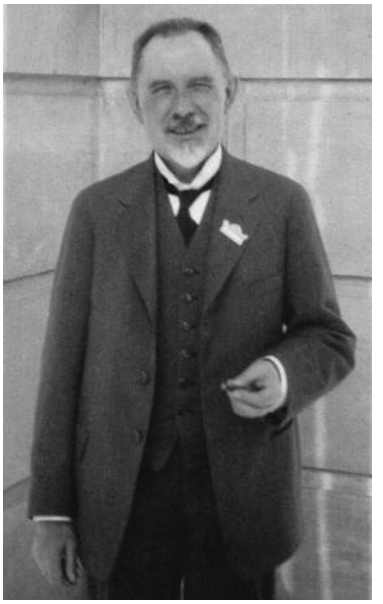
Figure 2.16 ▶
pH values for various fluids at 25°C. Lower values correspond to acidic fluids; higher values correspond to basic fluids.

▲ **pH strips.** The approximate pH of solutions can be determined in the lab by placing a drop on a pH strip. Various indicators are bound to a matrix that is affixed to a plastic strip. The indicators change color at different concentrations of $H^{\oplus}$, and the combination of various colors gives a more or less accurate reading of the pH. The strips shown here cover all pH readings from 0 to 14 but other pH strips can be used to cover narrower ranges.

BOX 2.4 THE LITTLE "p" IN pH

The term pH was first used in 1909 by Søren Peter Lauritz Sørensen, director of the Carlsberg Laboratories in Denmark. Sørensen never mentioned what the little "p" stood for (the "H" is obviously hydrogen). Many years later, some of the scientists who write chemistry textbooks began to associate the little "p" with the words *power* or *potential*. This association, as it turns out, is based on a rather tenuous connection in some of Sørensen's early papers. A recent investigation of the historical records by Jens G. Nøby suggests that the little "p" was an arbitrary choice based on Sørensen's use of p and q to stand for unknown variables in much the same way that we might use x and y today.

No matter what the historical origin, it's important to remember that the symbol pH now stands for the negative logarithm of the hydrogen ion concentration.



▲ **Søren Peter Lauritz Sørensen (1868–1939)**

Accurate measurements of pH are routinely made using a pH meter, an instrument that incorporates a selectively permeable glass electrode that is sensitive to $[H^{\oplus}]$. Measurement of pH sometimes facilitates the diagnosis of disease. The normal pH of human blood is 7.4—frequently referred to as physiological pH. The blood of patients suffering from certain diseases, such as diabetes, can have a lower pH, a condition called acidosis. The condition in which the pH of the blood is higher than 7.4, called alkalosis, can result from persistent, prolonged vomiting (loss of hydrochloric acid from the stomach) or from hyperventilation (excessive loss of carbonic acid as carbon dioxide).

KEY CONCEPT

pH is the negative logarithm of the proton ($H^{\oplus}$) concentration.

2.9 Acid Dissociation Constants of Weak Acids

Acids and bases that dissociate completely in water, such as hydrochloric acid and sodium hydroxide, are called strong acids and strong bases. Many other acids and bases, such as the amino acids from which proteins are made and the purines and pyrimidines from DNA and RNA, do not dissociate completely in water. These substances are known as weak acids and weak bases.

KEY CONCEPT

Weak acids and weak bases are compounds that only partially dissociate in water.

In order to understand the relationship between acids and bases let us consider the dissociation of HCl in water. Recall from Section 2.7 that we define an acid as a molecule that can donate a proton and a base as a proton acceptor. Acids and bases always come in pairs since for every proton donor there must be a proton acceptor. Both sides of the dissociation reaction will contain an acid and a base. Thus, the equilibrium reaction for the complete dissociation of HCl is

$$\underset{\text{acid}}{HCl} + \underset{\text{base}}{H_2O} \rightleftharpoons \underset{\text{base}}{Cl^{\ominus}} + \underset{\text{acid}}{H_3O^{\oplus}} \qquad (2.10)$$

HCl is an acid because it can donate a proton. In this case, the proton acceptor is water which is the base in this equilibrium reaction. On the other side of the equilibrium are $Cl^{\ominus}$ and the hydronium ion, $H_3O^{\oplus}$. The chloride ion is the base that corresponds to HCl after it has given up its proton. $Cl^{\ominus}$ is called the **conjugate base** of HCl which indicates that it is a base (i.e., can accept a proton) and is part of an acid–base pair (i.e., $HCl/Cl^{\ominus}$). Similarly, $H_3O^{\oplus}$ is the acid on the right-hand side of the equilibrium because it can donate a proton. $H_3O^{\oplus}$ is the **conjugate acid** of H_2O. Every base

has a corresponding conjugate acid and every acid has a corresponding conjugate base. Thus, HCl is the conjugate acid of $Cl^{\ominus}$ and H_2O is the conjugate base of $H_3O^{\oplus}$. Note that H_2O is the conjugate acid of $OH^{\ominus}$ if we are referring to the $H_2O/OH^{\ominus}$ acid–base pair.

In most cases throughout this book we will simplify reactions by ignoring the contribution of water and representing the hydronium ion as a simple proton.

$$HCl \rightleftharpoons H^{\oplus} + Cl^{\ominus} \tag{2.11}$$

This is a standard convention in biochemistry but, on the surface, it seems to violate the rule that both sides of the equilibrium reaction should contain a proton donor and a proton acceptor. Students should keep in mind that in such reactions the contributions of water molecules as proton acceptors and hydronium ions as the true proton donors are implied. In almost all cases we can safely ignore the contribution of water. This is the same principle that we applied to the reaction for the dissociation of water (Section 2.7) which we simplified by ignoring the contribution of one of the water molecules.

KEY CONCEPT

The contribution of water is implied in most acid/base dissociation reactions.

The reason why HCl is such a strong acid is because the equilibrium shown in Reaction 2.11 is shifted so far to the right that HCl is completely dissociated in water. In other words, HCl has a strong tendency to donate a proton when dissolved in water. This also means that the conjugate base, $Cl^{\ominus}$, is a very weak base because it will rarely accept a proton.

Acetic acid is the weak acid present in vinegar. The equilibrium reaction for the ionization of acetic acid is

$$\underset{\substack{\text{Acetic acid}\\\text{(weak acid)}}}{CH_3COOH} \overset{K_a}{\rightleftharpoons} H^{\oplus} + \underset{\substack{\text{Acetate anion}\\\text{(conjugate base)}}}{CH_3COO^{\ominus}} \tag{2.12}$$

We have left out the contribution of water molecules in order to simplify the reaction. We see that the acetate ion is the conjugate base of acetic acid. (We can also refer to acetic acid as the conjugate acid of the acetate ion.)

The equilibrium constant for the dissociation of a proton from an acid in water is called the **acid dissociation constant, K_a**. When the reaction reaches equilibrium, which happens very rapidly, the acid dissociation constant is equal to the concentration of the products divided by the concentration of the reactants. For Reaction 2.12 the acid dissociation constant is

$$K_a = \frac{[H^{\oplus}][CH_3COO^{\ominus}]}{[CH_3COOH]} \tag{2.13}$$

The K_a value for acetic acid at 25°C is 1.76×10^{-5} M. Because K_a values are numerically small and inconvenient in calculations it is useful to place them on a logarithmic scale. The parameter **pK_a** is defined by analogy with pH.

$$pK_a = -\log K_a = \log \frac{1}{K_a} \tag{2.14}$$

A pH value is a measure of the acidity of a solution and a pK_a value is a measure of the acid strength of a particular compound. The pK_a of acetic acid is 4.8.

When dealing with bases we need to consider their protonated forms in order to use Equation 2.13. These conjugate acids are very weak acids. In order to simplify calculations and make easy comparisons we measure the equilibrium constant (K_a) for the dissociation of a proton from the conjugate acid of a weak base. For example, the ammonium ion ($NH_4^{\oplus}$) can dissociate to form the base ammonia (NH_3) and $H^{\oplus}$.

$$NH_4^{\oplus} \rightleftharpoons NH_3 + H^{\oplus} \tag{2.15}$$

The acid dissociation constant (K_a) for this equilibrium is a measure of the strength of the base (ammonia, NH_3) in aqueous solution. The K_a values for several common substances are listed in Table 2.4.

Table 2.4 Dissociation constants and pK_a values of weak acids in aqueous solutions at 25°C

Acid	K_a(M)	pK_a
HCOOH (Formic acid)	1.77×10^{-4}	3.8
CH_3COOH (Acetic acid)	1.76×10^{-5}	4.8
$CH_3CHOHCOOH$ (Lactic acid)	1.37×10^{-4}	3.9
H_3PO_4 (Phosphoric acid)	7.52×10^{-3}	2.2
$H_2PO_4^{\ominus}$ (Dihydrogen phosphate ion)	6.23×10^{-8}	7.2
$HPO_4^{2\ominus}$ (Monohydrogen phosphate ion)	2.20×10^{-13}	12.7
H_2CO_3 (Carbonic acid)	4.30×10^{-7}	6.4
$HCO_3^{\ominus}$ (Bicarbonate ion)	5.61×10^{-11}	10.2
$NH_4^{\oplus}$ (Ammonium ion)	5.62×10^{-10}	9.2
$CH_3NH_3^{\oplus}$ (Methylammonium ion)	2.70×10^{-11}	10.7

From Equation 2.13 we see that the K_a for acetic acid is related to the concentration of $H^{\oplus}$ and to the ratio of the concentrations of the acetate ion and undissociated acetic acid. If we represent the conjugate acid as HA and the conjugate base as $A^{\ominus}$ then taking the logarithm of such equations gives the general equation for any acid–base pair.

$$HA \rightleftharpoons H^{\oplus} + A^{\ominus} \qquad \log K_a = \log \frac{[H^{\oplus}][A^{\ominus}]}{[HA]} \tag{2.16}$$

Since $\log(xy) = \log x + \log y$, Equation 2.16 can be rewritten as

$$\log K_a = \log[H^{\oplus}] + \log \frac{[A^{\ominus}]}{[HA]} \tag{2.17}$$

Rearranging Equation 2.17 gives

$$-\log[H^{\oplus}] = -\log K_a + \log \frac{[A^{\ominus}]}{[HA]} \tag{2.18}$$

KEY CONCEPT

The pH of a solution of a weak acid or base at equilibrium can be calculated by combining the pK_a of the ionization reaction and the final concentrations of the proton acceptor and proton donor species.

The negative logarithms in Equation 2.18 have already been defined as pH and pK_a (Equations 2.9 and 2.14, respectively). Thus,

$$pH = pK_a + \log \frac{[A^{\ominus}]}{[HA]} \tag{2.19}$$

or

$$pH = pK_a + \log \frac{[\text{Proton acceptor}]}{[\text{Proton donor}]} \tag{2.20}$$

Equation 2.20 is one version of the **Henderson–Hasselbalch equation**. It defines the pH of a solution in terms of the pK_a of the weak acid form of the acid–base pair and the logarithm of the ratio of concentrations of the dissociated species (conjugate base) to the protonated species (weak acid). Note that the greater the concentration of the proton acceptor (conjugate base) relative to that of the proton donor (weak acid), the lower the concentration of $H^{\oplus}$ and the higher the pH. (Remember that pH is the *negative* log of $H^{\oplus}$ concentration. A high concentration of $H^{\oplus}$ means low pH.) This

makes intuitive sense since the concentration of $A^{\ominus}$ is identical to the concentration of $H^{\oplus}$ in simple dissociation reactions. If more HA dissociates the concentration of $A^{\ominus}$ will be higher and so will the concentration of $H^{\oplus}$. When the concentrations of a weak acid and its conjugate base are exactly the same the pH of the solution is equal to the pK_a of the acid (since the ratio of concentrations equals 1.0, and the logarithm of 1.0 equals zero).

The Henderson–Hasselbalch equation is used to determine the final pH of a weak acid solution once the dissociation reaction reaches equilibrium as illustrated in Sample Calculation 2.1 for acetic acid. These calculations are more complicated than those involving strong acids such as HCl. As noted in Section 2.8, the pH of an HCl solution is easily determined from the amount of HCl that is present since the final concentration of $H^{\oplus}$ is equal to the initial concentration of HCl when the solution is made up. In contrast, weak acids are only partially dissociated in water so it makes sense that the pH depends on the acid dissociation constant. The pH decreases (more $H^{\oplus}$) as more weak acid is added to water but the increase in $H^{\oplus}$ is not linear with initial HA concentration. This is because the numerator in Equation 2.16 is the *product* of the $H^{\oplus}$ and $A^{\ominus}$ concentrations.

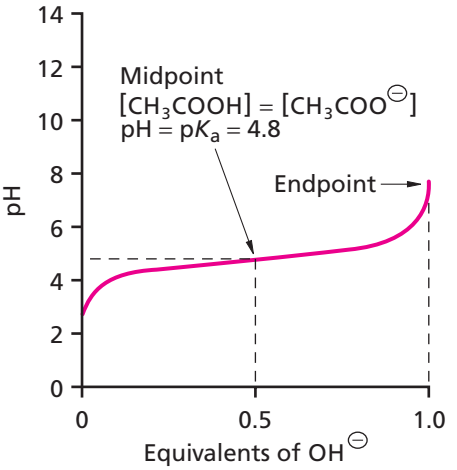
The Henderson–Hasselbalch equation applies to other acid–base combinations as well and not just to those involving weak acids. When dealing with a weak base, for example, the numerator and denominator of Equation 2.20 become [weak base] and [conjugate acid], respectively. The important point to remember is that the equation refers to the concentration of the proton acceptor divided by the concentration of the proton donor.

The pK_a values of weak acids are determined by titration. Figure 2.17 shows the titration curve for acetic acid. In this example, a solution of acetic acid is titrated by adding small aliquots of a strong base of known concentration. The pH of the solution is measured and plotted versus the number of molar equivalents of strong base added during the titration. Note that since acetic acid has only one ionizable group (its carboxyl group) only one equivalent of a strong base is needed to completely titrate acetic acid to its conjugate base, the acetate anion. When the acid has been titrated with one-half an equivalent of base the concentration of undissociated acetic acid exactly equals the concentration of the acetate anion. The resulting pH, 4.8, is thus the experimentally determined pK_a for acetic acid.

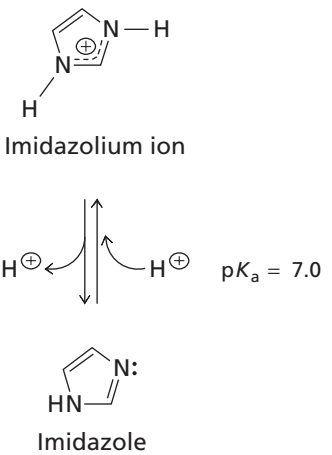
Constructing an ideal titration curve is a useful exercise for reinforcing the relationship between pH and the ionization state of a weak acid. You can use the Henderson–Hasselbalch equation to calculate the pH that results from adding increasing amounts of a strong base such as NaOH to a weak acid such as the imidazolium ion $pK_a = 7.0$. Adding base converts the imidazolium ion to its conjugate base, imidazole (Figure 2.18). The shape of the titration curve is easy to visualize if you calculate the pH when the ratio of conjugate base to acid is 0.01, 0.1, 1, 10, and 100. Calculate pH values at other ratios until you are satisfied that the curve is relatively flat near the midpoint and steeper at the ends.

Similarly shaped titration curves can be obtained for each of the five monoprotic acids (acids having only one ionizable group) listed in Table 2.4. All would exhibit the same general shape as Figure 2.17 but the inflection point representing the midpoint of titration (one-half an equivalent titrated) would fall lower on the pH scale for a stronger acid (such as formic acid or lactic acid) and higher for a weaker acid (such as ammonium ion or methylammonium ion).

Titration curves of weak acids illustrate a second important use of the Henderson–Hasselbalch equation. In this case, the final pH is the result of mixing the weak acid (HA) and a strong base ($OH^{\ominus}$). The base combines with $H^{\oplus}$ ions to form water molecules, H_2O. This reduces the concentration of $H^{\oplus}$ and raises the pH. As the titration of the weak acid proceeds it dissociates in order to restore its equilibrium with $OH^{\ominus}$ and H_2O. The net result is that the final concentration of $A^{\ominus}$ is much higher, and the concentration of HA is much lower, than when we are dealing with the simple case where the pH is determined only by the dissociation of the weak acid in water (i.e., a solution of HA in H_2O).

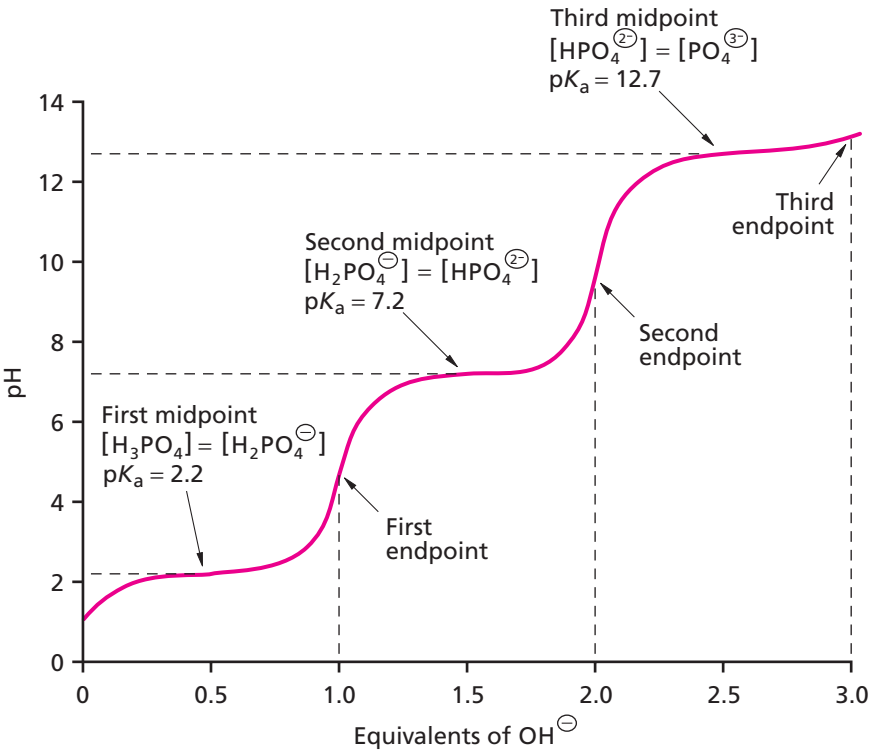


▲ **Figure 2.17**
Titration of acetic acid (CH_3COOH) with aqueous base ($OH^{\ominus}$). There is an inflection point (a point of minimum slope) at the midpoint of the titration, when 0.5 equivalent of base has been added to the solution of acetic acid. This is the point at which $[CH_3COOH] = [CH_3COO^{\ominus}]$ and $pH = pK_a$. The pK_a of acetic acid is thus 4.8. At the endpoint, all the molecules of acetic acid have been titrated to the conjugate base, acetate.



▲ **Figure 2.18**
Titration of the imidazolium ion.

Figure 2.19 ▶
Titration curve for H_3PO_4. Three inflection points (at 0.5, 1.5, and 2.5 equivalents of strong base added) correspond to the three pK_a values for phosphoric acid (2.2, 7.2, and 12.7).

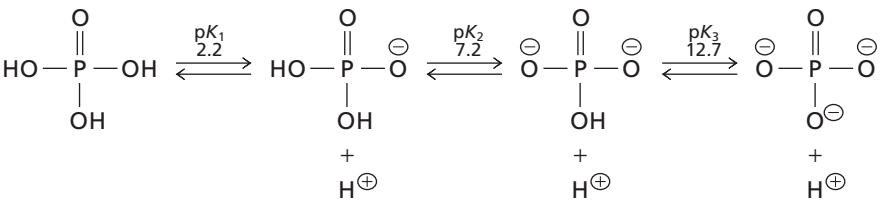


▲ Cola beverages contain phosphoric acid in order to make the drink more acidic. The concentration of phosphoric acid is about 1 mM. This concentration should make the pH about 3 in the absence of any other ingredients that may contribute to acidity.

Phosphoric acid (H_3PO_4) is a polyprotic acid. It contains three different hydrogen atoms that can dissociate to form $H^{\oplus}$ ions and corresponding conjugate bases with one, two, or three negative charges. The dissociation of the first proton occurs readily and is associated with a large acid dissociation constant of 7.53×10^{-3} M and a pK_a of 2.2 in aqueous solution. The dissociations of the second and third protons occur progressively less readily because they have to dissociate from a molecule that is already negatively charged.

Phosphoric acid requires three equivalents of strong base for complete titration and three pK_a values are evident from its titration curve (Figure 2.19). The three pK_a values reflect the three equilibrium constants and thus the existence of four possible ionic species (conjugate acids and bases) of inorganic phosphate. At physiological pH (7.4) the predominant species of inorganic phosphate are $H_2PO_4^{\ominus}$ and $HPO_4^{2\ominus}$. At pH 7.2 these two species exist in equal concentrations. The concentrations of H_3PO_4 and $PO_4^{3\ominus}$ are so low at pH 7.4 that they can be ignored. This is generally the case for a minor species when the pH is more than two units away from its pK_a.

(2.21)



Many biologically important acids and bases, including the amino acids described in Chapter 3, have two or more ionizable groups. The number of pK_a values for such substances is equal to the number of ionizable groups. The pK_a values can be experimentally determined by titration.

Sample Calculation 2.1 CALCULATING THE pH OF WEAK ACID SOLUTIONS

Q: What is the pH of a solution of 0.1 M acetic acid?

A: The acid dissociation constant of acetic acid is 1.76×10^{-5} M. Acetic acid dissociates in water to form acetate and $H^{\oplus}$. We need to determine $[H^{\oplus}]$ when the reaction reaches equilibrium.

Let the final $H^{\oplus}$ concentration be represented by the unknown quantity x. At equilibrium the concentration of acetate ion will also be x and the final concentration of acetic acid will be $[0.1\ M - x]$. Thus,

$$1.76 \times 10^{-5} = \frac{[H^{\oplus}][CH_3COO^{\ominus}]}{[CH_3COOH]} = \frac{x^2}{(0.1 - x)}$$

rearranging gives

$$1.76 \times 10^{-6} - 1.76 \times 10^{-5}x = x^2$$
$$x^2 + 1.76 \times 10^{-5}\,x - 1.76 \times 10^{-6} = 0$$

This equation is a typical quadratic equation of the form $ax^2 + bx + c = 0$, where $a = 1$, $b = 1.76 \times 10^{-5}$, and $c = -1.76 \times 10^{-6}$. Solve for x using the standard formula

$$x = \frac{-b \pm \sqrt{(b^2 - 4ac)}}{2a}$$

$$= \frac{-1.76 \times 10^{-5} \pm \sqrt{((1.76 \times 10^{-5})^2 - 4(1.76 \times 10^{-6}))}}{2}$$

$$x = 0.00132 \quad \text{or} \quad -0.00135 \text{ (reject the negative answer)}$$

The hydrogen ion concentration is 0.00132 M and the pH is

$$pH = -\log[H^{\oplus}] = -\log(0.00132) = -(-2.88) = 2.9$$

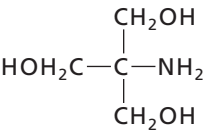
Note that the contribution of hydrogen ions from the dissociation of water (10^{-7}) is several orders of magnitude lower than the concentration of hydrogen ions from acetic acid. It is standard practice to ignore the ionization of water in most calculations as long as the initial concentration of weak acid is greater than 0.001 M.

The amount of acetic acid that dissociates to form $H^{\oplus}$ and $CH_3COO^{\ominus}$ is 0.0013 M when the initial concentration is 0.1 M. This means that only 1.3% of the acetic acid molecules dissociate and the final concentration of acetic acid $([CH_3COOH])$ is 98.7% of the initial concentration. In general, the percent dissociation of dilute solutions of weak acids is less than 10% and it is a reasonable approximation to assume that the final concentration of the acid form is the same as its initial concentration. This approximation has very little effect on the calculated pH and it has the advantage of avoiding quadratic equations.

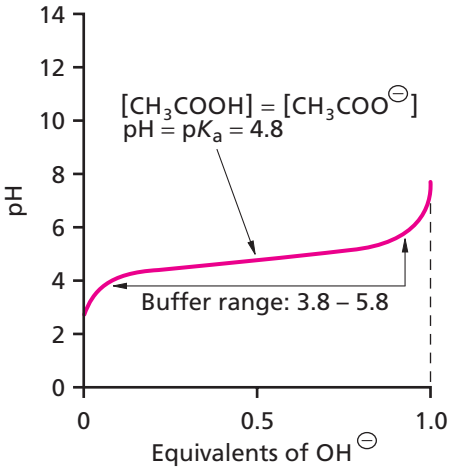
Assuming that the concentration of CH_3COOH at equilibrium is 0.1 M and the concentration of $H^{\oplus}$ is x,

$$K_a = 1.76 \times 10^{-5} = \frac{x^2}{0.1} \qquad x = 1.33 \times 10^{-3}$$

$$pH = -\log(1.33 \times 10^{-3}) = 2.88 = 2.9$$



▲ **Tris buffers.** Tris, or tris (hydroxymethyl) aminomethane, is a common buffer in biochemistry labs. Its pK_a of 8.06 makes it ideal for preparation of buffers in the physiological range.



▲ **Figure 2.20**
Buffer range of acetic acid. For CH_3COOH + $CH_3COO^{\ominus}$ the pK_a is 4.8 and the most effective buffer range is from pH 3.8 to pH 5.8.

2.10 Buffered Solutions Resist Changes in pH

If the pH of a solution remains nearly constant when small amounts of strong acid or strong base are added the solution is said to be **buffered**. The ability of a solution to resist changes in pH is known as its buffer capacity. Inspection of the titration curves of acetic acid (Figure 2.17) and phosphoric acid (Figure 2.19) reveals that the most effective buffering, indicated by the region of minimum slope on the curve, occurs when the concentrations of a weak acid and its conjugate base are equal—in other words, when the pH equals the pK_a. The effective range of buffering by a mixture of a weak acid and its conjugate base is usually considered to be from one pH unit below to one pH unit above the pK_a.

Most *in vitro* biochemical experiments involving purified molecules, cell extracts, or intact cells are performed in the presence of a suitable buffer to ensure a stable pH. A number of synthetic compounds with a variety of pK_a values are often used to prepare buffered solutions but naturally occurring compounds can also be used as buffers. For example, mixtures of acetic acid and sodium acetate (pK_a= 4.8) can be used for the pH range from 4 to 6 (Figure 2.20) and mixtures of KH_2PO_4 and K_2HPO_4 (pK_a = 7.2) can be used in the range from 6 to 8. The amino acid glycine (pK_a = 9.8) is often used in the range from 9 to 11.

When preparing buffers the acid solution (e.g., acetic acid) supplies the protons and some of the protons are taken up by combining with the conjugate base (e.g., acetate). The conjugate base is added as a solution of a salt (e.g., sodium acetate). The salt dissociates completely in solution providing free conjugate base and no protons. Sample Calculation 2.2 illustrates one way to prepare a buffer solution.

Sample Calculation 2.2 BUFFER PREPARATION

Q: Acetic acid has a pK_a of 4.8. How many milliliters of 0.1 M acetic acid and 0.1 M sodium acetate are required to prepare 1 liter of 0.1 M buffer solution having a pH of 5.8?

A: Substitute the values for the pK_a and the desired pH into the Henderson–Hasselbalch equation (Equation 2.20).

$$5.8 = 4.8 + \log\frac{[\text{Acetate}]}{[\text{Acetic acid}]}$$

Solve for the ratio of acetate to acetic acid.

$$\log\frac{[\text{Acetate}]}{[\text{Acetic acid}]} = 5.8 - 4.8 = 1.0$$

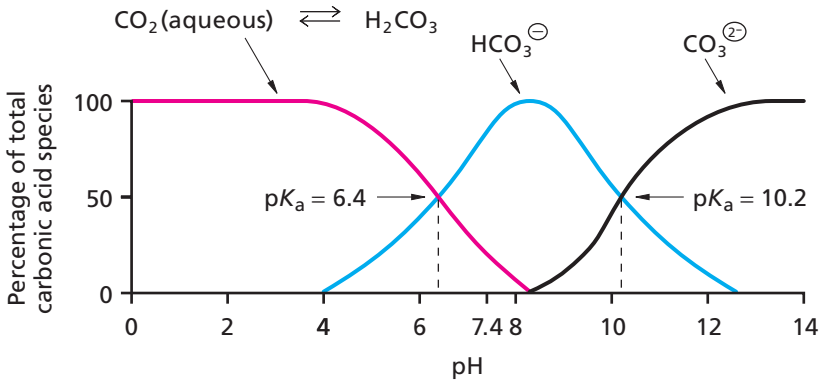
$$[\text{Acetate}] = 10\,[\text{Acetic acid}]$$

For each volume of acetic acid, 10 volumes of acetate must be added (making a total of 11 volumes of the two ionic species). Multiply the proportion of each component by the desired volume.

$$\text{Acetic acid needed:} \quad \frac{1}{11} \times 1000 \text{ ml} = 91 \text{ ml}$$

$$\text{Acetate needed:} \quad \frac{10}{11} \times 1000 \text{ ml} = 909 \text{ ml}$$

Note that when the ratio of [conjugate base] to [conjugate acid] is 10:1, the pH is exactly one unit above the pK_a. If the ratio were. 1:10, the pH would be one unit below the pK_a.



▶ **Figure 2.21**
Percentages of carbonic acid and its conjugate base as a function of pH. In an aqueous solution at pH 7.4 (the pH of blood) the concentrations of carbonic acid (H_2CO_3) and bicarbonate ($HCO_3^{\ominus}$) are substantial, but the concentration of carbonate ($CO_3^{2\ominus}$) is negligible.

An excellent example of buffer capacity is found in the blood plasma of mammals, which has a remarkably constant pH. Consider the results of an experiment that compares the addition of an aliquot of strong acid to a volume of blood plasma with a similar addition of strong acid to either physiological saline (0.15 M NaCl) or water. When 1 milliliter of 10 M HCl (hydrochloric acid) is added to 1 liter of physiological saline or water that is initially at pH 7.0 the pH is lowered to 2.0 (in other words, $[H^{\oplus}]$ from HCl is diluted to 10^{-2} M). However, when 1 milliliter of 10 M HCl is added to 1 liter of human blood plasma at pH 7.4 the pH is lowered to only 7.2—impressive evidence for the effectiveness of physiological buffering.

The pH of blood is primarily regulated by the carbon dioxide–carbonic acid–bicarbonate buffer system. A plot of the percentages of carbonic acid (H_2CO_3) and its conjugate base as a function of pH is shown in Figure 2.21. Note that the major components at pH 7.4 are carbonic acid and the bicarbonate anion ($HCO_3^{\ominus}$).

The buffer capacity of blood depends on equilibria between gaseous carbon dioxide (which is present in the air spaces of the lungs), aqueous carbon dioxide (which is produced by respiring tissues and dissolved in blood), carbonic acid, and bicarbonate. As shown in Figure 2.21, the equilibrium between bicarbonate and its conjugate base, carbonate ($CO_3^{2\ominus}$), does not contribute significantly to the buffer capacity of blood because the pK_a of bicarbonate is 10.2—too far from physiological pH to have an effect on the buffering of blood.

The first of the three relevant equilibria of the carbon dioxide–carbonic acid–bicarbonate buffer system is the dissociation of carbonic acid to bicarbonate.

$$H_2CO_3 \rightleftharpoons H^{\oplus} + HCO_3^{\ominus} \qquad \textbf{(2.22)}$$

This equilibrium is affected by a second equilibrium in which dissolved carbon dioxide is in equilibrium with its hydrated form, carbonic acid.

$$CO_2(\text{aqueous}) + H_2O \rightleftharpoons H_2CO_3 \qquad \textbf{(2.23)}$$

These two reactions can be combined into a single equilibrium reaction where the acid is represented as CO_2 dissolved in water:

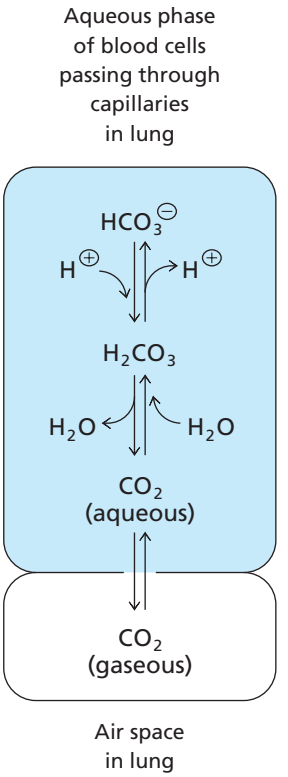
$$CO_2(\text{aqueous}) + H_2O \rightleftharpoons H^{\oplus} + HCO_3^{\ominus} \qquad \textbf{(2.24)}$$

The pK_a of the acid is 6.4.

Finally, CO_2 (gaseous) is in equilibrium with CO_2 (aqueous).

$$CO_2(\text{gaseous}) \rightleftharpoons CO_2(\text{aqueous}) \qquad \textbf{(2.25)}$$

The regulation of the pH of blood afforded by these three equilibria is shown schematically in Figure 2.22. When the pH of blood falls due to a metabolic process that produces excess $H^{\oplus}$ the concentration of H_2CO_3 increases momentarily but H_2CO_3



▲ **Figure 2.22**
Regulation of the pH of blood in mammals. The pH of blood is controlled by the ratio of $[HCO_3^{\ominus}]$ to pCO_2 in the air spaces of the lungs. When the pH of blood decreases due to excess $H^{\oplus}$, pCO_2 increases in the lungs, restoring the equilibrium. When the concentration of $HCO_3^{\ominus}$ rises because the pH of blood increases, CO_2 (gaseous) dissolves in the blood, again restoring the equilibrium.

rapidly loses water to form dissolved CO_2 (aqueous) which enters the gaseous phase in the lungs and is expired as CO_2 (gaseous). An increase in the partial pressure of CO_2 (pCO_2) in the air expired from the lungs thus compensates for the increased hydrogen ions. Conversely, when the pH of the blood rises the concentration of $HCO_3^{\ominus}$ increases transiently but the pH is rapidly restored as the breathing rate changes and the CO_2 (gaseous) in the lungs is converted to CO_2 (aqueous) and then to H_2CO_3 in the capillaries of the lungs. Again, the equilibrium of the blood buffer system is rapidly restored by changing the partial pressure of CO_2 in the lungs.

Within cells, both proteins and inorganic phosphate contribute to intracellular buffering. Hemoglobin is the strongest buffer in blood cells other than the carbon dioxide–carbonic acid–bicarbonate buffer. As mentioned earlier, the major species of inorganic phosphate present at physiological pH are $H_2PO_4^{\ominus}$ and $HPO_4^{2\ominus}$ reflecting the second pK_a (pK_2) value for phosphoric acid, 7.2.

Summary

1. The water molecule has a permanent dipole because of the uneven distribution of charge in O—H bonds and their angled arrangement.
2. Water molecules can form hydrogen bonds with each other. Hydrogen bonding contributes to the high specific heat and heat of vaporization of water.
3. Because it is polar, water can dissolve ions. Water molecules form a solvation sphere around each dissolved ion. Organic molecules may be soluble in water if they contain ionic or polar functional groups that can form hydrogen bonds with water molecules.
4. The hydrophobic effect is the exclusion of nonpolar substances by water molecules. Detergents, which contain both hydrophobic and hydrophilic portions, form micelles when suspended in water; these micelles can trap insoluble substances in a hydrophobic interior. Chaotropes enhance the solubility of nonpolar compounds in water.
5. The major noncovalent interactions that determine the structure and function of biomolecules are electrostatic interactions and hydrophobic interactions. Electrostatic interactions include charge–charge interactions, hydrogen bonds, and van der Waals forces.
6. Under cellular conditions, macromolecules do not spontaneously hydrolyze, despite the presence of high concentrations of water. Specific enzymes catalyze their hydrolysis, and other enzymes catalyze their energy-requiring biosynthesis.
7. At 25°C, the product of the proton concentration ($[H^{\oplus}]$) and the hydroxide concentration ($[OH^{\ominus}]$) is 1.0×10^{-14} M^2, a constant designated K_w (the ion-product constant for water). Pure water ionizes to produce 10^{-7} M $H^{\oplus}$ and 10^{-7} M $OH^{\ominus}$.
8. The acidity or basicity of an aqueous solution depends on the concentration of $H^{\oplus}$ and is described by a pH value, where pH is the negative logarithm of the hydrogen ion concentration.
9. The strength of a weak acid is indicated by its pK_a value. The Henderson–Hasselbalch equation defines the pH of a solution of weak acid in terms of the pK_a and the concentrations of the weak acid and its conjugate base.
10. Buffered solutions resist changes in pH. In human blood, a constant pH of 7.4 is maintained by the carbon dioxide–carbonic acid–bicarbonate buffer system.

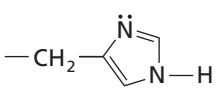
Problems

1. The side chains of some amino acids possess functional groups that readily form hydrogen bonds in aqueous solution. Draw the hydrogen bonds likely to form between water and the following amino acid side chains:

   (a) CH_2OH

   (b) $CH_2C(O)NH_2$

   (c)

2. State whether each of the following compounds is polar, whether it is amphipathic, and whether it readily dissolves in water.

   (a) $HO-CH_2-\underset{\displaystyle OH}{\underset{|}{CH}}-CH_2-OH$
   Glycerol

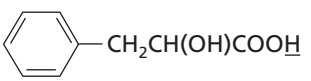
   (b) $CH_3(CH_2)_{14}-CH_2-OPO_3^{2\ominus}$
   Hexadecanyl phosphate

   (c) $CH_3-(CH_2)_{10}-COO^{\ominus}$
   Laurate

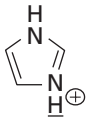
   (d) $H_3\overset{\oplus}{N}-CH_2-COO^{\ominus}$
   Glycine

3. Osmotic lysis is a gentle method of breaking open animal cells to free intracellular proteins. In this technique, cells are suspended in a solution that has a total molar concentration of solutes much less than that found naturally inside cells. Explain why this technique might cause cells to burst.
4. Each of the following molecules is dissolved in buffered solutions of: (a) pH = 2 and (b) pH = 11. For each molecule, indicate the solution in which the charged species will predominate. (Assume that the added molecules do not appreciably change the pH of the solution.)

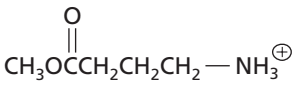
   (a) Phenyl lactic acid pK_a = 4

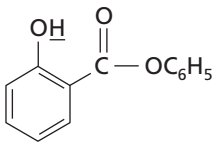
(b) Imidazole pK_a = 7



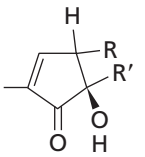
(c) *O*-methyl-γ-aminobutyrate pK_a = 9.5



(d) Phenyl salicylate pK_a = 9.6

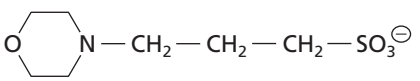


5. Use Figure 2.16 to determine the concentration of $H^{\oplus}$ and $OH^{\ominus}$ in:
   (a) tomato juice
   (b) human blood plasma
   (c) 1 M ammonia
6. The interaction between two (or more) molecules in solution can be mediated by specific hydrogen bond interactions. Phorbol esters can act as a tumor promoter by binding to certain amino acids that are part of the enzyme protein kinase C (PKC). Draw the hydrogen bonds expected in the complex formed between the tumor promoter phorbol and the glycine portion of PKC: —$NHCH_2C(O)$—

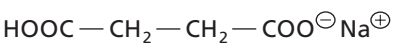


Phorbol

7. What is the concentration of a lactic acid buffer (pK_a = 3.9) that contains 0.25 M $CH_3CH(OH)COOH$ and 0.15 M $CH_3CH(OH)COO^{\ominus}$? What is the pH of this buffer?
8. You are instructed to prepare 100 ml of a 0.02 M sodium phosphate buffer, pH 7.2, by mixing 50 ml of solution A (0.02M Na_2HPO_4) and 50 ml of solution B (0.02 M NaH_2PO_4). Refer to Table 2.4 to explain why this procedure provides an effective buffer at the desired pH and concentration.
9. What are the effective buffering ranges of MOPS (3-(*N*-morpholino)propanesulfonic acid) and SHS (sodium hydrogen succinate)?



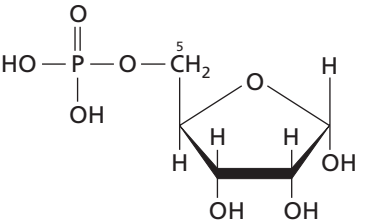
MOPS



SHS

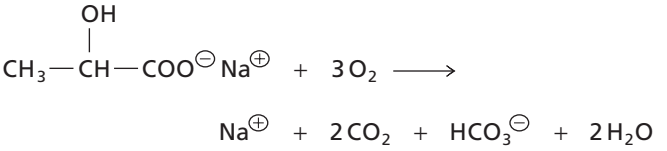
The nitrogen atom of MOPS can be protonated (pK_a = 7.2). The carboxyl group of SHS can be ionized (pK_a = 5.5). Calculate the ratio of basic to acidic species for each buffer at pH 6.5.

10. Many phosphorylated sugars (phosphate esters of sugars) are metabolic intermediates. The two ionizable —OH groups of the phosphate group of the monophosphate ester of ribose (ribose 5-phosphate) have pK_a values of 1.2 and 6.6. The fully protonated form of α-D-ribose 5-phosphate has the structure shown below.

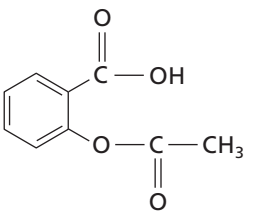


α-D-Ribose 5-phosphate

   (a) Draw, in order, the ionic species formed upon titration of this phosphorylated sugar from pH 0.0 to pH 10.0.
   (b) Sketch the titration curve for ribose 5-phosphate.
11. Normally, gaseous CO_2 is efficiently expired in the lungs. Under certain conditions, such as obstructive lung disease or emphysema, expiration is impaired. The resulting excess of CO_2 in the body may lead to respiratory acidosis, a condition in which excess acid accumulates in bodily fluids. How does excess CO_2 lead to respiratory acidosis?
12. Organic compounds in the diets of animals are a source of basic ions and may help combat nonrespiratory types of acidosis. Many fruits and vegetables contain salts of organic acids that can be metabolized, as shown below for sodium lactate. Explain how the salts of dietary acids may help alleviate metabolic acidosis.



13. Absorption of food in the stomach and intestine depends on the ability of molecules to penetrate the cell membranes and pass into the bloodstream. Because hydrophobic molecules are more likely to be absorbed than hydrophilic or charged molecules, the absorption of orally administered drugs may depend on their pK_a values and the pH in the digestive organs. Aspirin (acetylsalicylic acid) has an ionizable carboxyl group (pK_a = 3.5). Calculate the percentage of the protonated form of aspirin available for absorption in the stomach (pH = 2.0) and in the intestine (pH = 5.0).

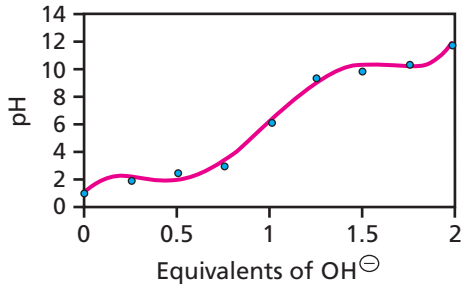


Aspirin

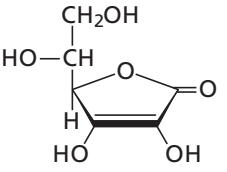
14. What percent of glycinamide, $^{\oplus}H_3NCH_2CONH_2$ (pK_a = 8.20) is unprotonated at (a) pH 7.5, (b) pH 8.2, and (c) pH 9.0?

15. Refer to the following table and titration curve to determine which compound from the table is illustrated by the titration curve.

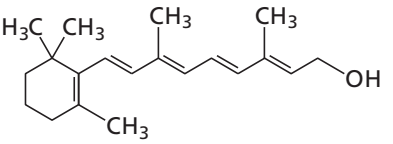
Compound	pK_1	pK_2	pK_3
Phosphoric acid	2.15	7.20	12.15
Acetic acid	4.76		
Succinic acid	4.21	5.64	
Boric acid	9.24	12.74	
Glycine	2.40	9.80	



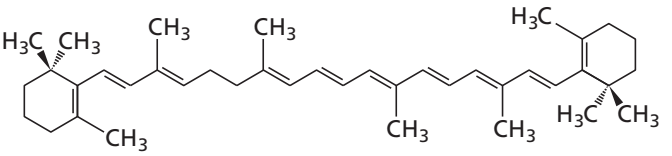
16. Predict which of the following substances are soluble in water.



(a) Vitamin C



(b) Vitamin A



(c) β-carotene

17. The ion product for water at 0°C is 1.14×10^{-15}, and at 100°C it is about 4.0×10^{-13}. What is the actual neutral pH for extremophiles living at 0°C and 100°C?

18. What is the approximate pH of a solution of 6 M HCl? Why doesn't the scale in Figure 2.16 accommodate the pH of such a solution?

Selected Readings

Water

Chaplin, M. F. (2001). Water, its importance to life. *Biochem. and Mol. Biol. Education* 29:54–59.

Dix, J. A. and Verkman, A. S. (2008). Crowding effects on diffusion in solutions and cells. *Annu. Rev. Biophys.* 37:247–263.

Stillinger, F. H. (1980). Water revisited. *Science* 209:451–457.

Verkman, A. S. (2001). Solute and macromolecular diffusion in cellular aqueous compartments. *Trends Biochem Sci.* 27:27–33.

Noncovalent Interactions

Fersht, A. R. (1987). The hydrogen bond in molecular recognition. *Trends Biochem. Sci.* 12:301–304.

Frieden, E. (1975). Non-covalent interactions. *J. Chem. Educ.* 52:754–761.

Tanford, C. (1980). *The Hydrophobic Effect: Formation of Micelles and Biological Membranes,* 2nd ed. (New York: John Wiley & Sons).

Biochemical Calculations

Montgomery, R., and Swenson, C. A. (1976). *Quantitative Problems in Biochemical Sciences,* 2nd ed. (San Francisco: W. H. Freeman).

Segel, I. H. (1976). *Biochemical Calculations: How to Solve Mathematical Problems in General Biochemistry*, 2nd ed. (New York: John Wiley & Sons).

pH and Buffers

Stoll, V. S., and Blanchard, J. S. (1990). Buffers: principles and practice. *Methods Enzymol.* 182:24–38.

Nørby, J. G. (2000). The origin and meaning of the little p in pH. *Trends Biochem. Sci.* 25:36–37.