

CHAPTER 3

The Chemistry of Life: Organic Compounds

Both inorganic and organic forms of carbon occur widely in nature. The human, plant, paper, and the wood of the desk in the photograph all contain **organic compounds**, formed from backbones of covalently bonded carbon. Inorganic elemental carbon is represented by the diamond and the pencil “lead,” which is called graphite. Some very simple carbon compounds are also considered inorganic, especially if the carbon is not bonded to hydrogen. Organic compounds are so named because at one time they were thought to be produced only by living (organic) organisms. In 1928 the German chemist Friedrich Wöhler synthesized urea, a metabolic waste product. Since that time, scientists have learned to synthesize many organic molecules and have discovered organic compounds not found in organisms.

In this chapter we focus on some of the major groups of organic compounds important in organisms, including carbohydrates, lipids, proteins, and nucleic acids (DNA and RNA). Organic compounds are the main structural components of cells and tissues. They participate in and regulate metabolic reactions, transmit information, and provide energy for life processes. Evolution involves chemical changes in the organic compounds produced by organisms.

More than five million organic compounds have been identified. Perhaps because it can form a greater variety of molecules than any other element, carbon is the central component of all organic compounds. There are many reasons that organic molecules are so diverse. **Hydrocarbons**, organic compounds consisting only of carbon and hydrogen, can be produced in a wide variety of three-dimensional shapes. Furthermore, the carbon atom can form bonds with a greater number of different elements than any other type of atom. The addi-



(Kenneth Knott/Fine Light Photography)

tion of chemical groups containing atoms of other elements, especially oxygen, nitrogen, phosphorus, and sulfur, can profoundly change the properties of an organic molecule.

Further diversity is provided by the fact that a great many organic compounds found in organisms are extremely large **macromolecules**. Cells construct these from simpler modular subunits. For example, protein molecules are built from smaller compounds called amino acids.

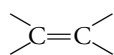
LEARNING OBJECTIVES

AFTER YOU HAVE STUDIED THIS CHAPTER YOU SHOULD BE ABLE TO

1. Distinguish between organic and inorganic compounds.
2. Describe the properties of carbon that make it the central component of organic compounds.
3. Distinguish among the three principal types of isomers.
4. Identify the major functional groups present in organic compounds and describe their properties.
5. Compare the functions and chemical compositions of the major groups of organic compounds: carbohydrates, lipids, proteins, and nucleic acids.
6. Distinguish among monosaccharides, disaccharides, and polysaccharides. Compare storage polysaccharides with structural polysaccharides.
7. Distinguish among neutral fats, phospholipids, and steroids; describe the compositions, characteristics, and biological functions of each.
8. Sketch the structure of an amino acid. Explain how amino acids are grouped into classes based on the characteristics of their side chains.
9. Distinguish among the levels of organization of protein molecules.
10. Describe the components of a nucleotide. Name some nucleic acids and discuss the importance of these compounds in living organisms.

CARBON ATOMS FORM AN ENORMOUS VARIETY OF STRUCTURES

Carbon has unique properties that permit formation of the carbon backbones of the large, complex molecules essential to life. Because a carbon atom has four valence electrons, it can complete its valence shell by forming a total of four covalent bonds. Each bond can link it to another carbon atom or to an atom of a different element. Carbon is particularly well suited to serve as the backbone of a large molecule because carbon-to-carbon bonds are strong and not easily broken. However, they are not so strong that it would be impossible for cells to break them. Carbon-to-carbon bonds are not limited to single bonds (based on sharing of one electron pair). Two carbon atoms can share two electron pairs with each other, forming double bonds:



In some compounds, triple carbon-to-carbon bonds are formed:



Carbon chains can be unbranched or branched, and carbon atoms can also be joined into rings (Fig. 3–1). Rings and chains are joined in some compounds.

The molecules in the cell are analogous to the components of a machine. Each component has a shape that permits it to fill certain roles and to interact with other components (often with a complementary shape). Similarly, the shape of a molecule is important in determining its biological properties and function. Carbon atoms are able to link to each other and to other atoms to produce a wide variety of three-dimensional molecular shapes. This is because the four covalent bonds of carbon do not form in a single plane. Instead, as discussed in Chapter 2, the valence electron orbitals become elongated and project from the carbon atom toward the corners of a tetrahedron (Fig. 3–2). The structure is highly symmetrical, with an angle of about 109.5 degrees between any two of these

bonds. Keep in mind that, for simplicity, many of the figures in this book are drawn as two-dimensional graphic representations of three-dimensional molecules. For example, hydrocarbon chains, such as those seen in Figure 3–1, are not actually straight but have a three-dimensional zigzag structure.

Generally, there is freedom of rotation around each carbon-to-carbon single bond. This property permits organic molecules to be flexible and to assume a variety of shapes, depending on the extent to which each single bond is rotated. Double and triple bonds do not permit rotation, so regions of a molecule with such bonds tend to be inflexible.

One reason for the very great number of possible carbon-containing compounds is the fact that the same components usually can link together in more than one pattern, generating an even wider variety of molecular shapes.

ISOMERS HAVE THE SAME MOLECULAR FORMULA, BUT DIFFERENT STRUCTURES

Compounds with the same molecular formulas but different structures and thus different properties are called **isomers**. Isomers do not have identical physical or chemical properties and may have different common names. Cells can distinguish between isomers. Usually, one isomer is biologically active and the other is not. Three types of isomers are structural isomers, geometric isomers, and enantiomers.

Structural isomers are compounds that differ in the covalent arrangements of their atoms. For example, Figure 3–3*a* illustrates two structural isomers with the molecular formula $\text{C}_2\text{H}_6\text{O}$. Similarly, there are two structural isomers of the four-carbon hydrocarbon butane (C_4H_{10}), one with a straight chain and the other with a branched chain (isobutane). Large compounds have more possible structural isomers. There are only two structural isomers of butane, but there can be up to 366,319 isomers of $\text{C}_{20}\text{H}_{42}$.

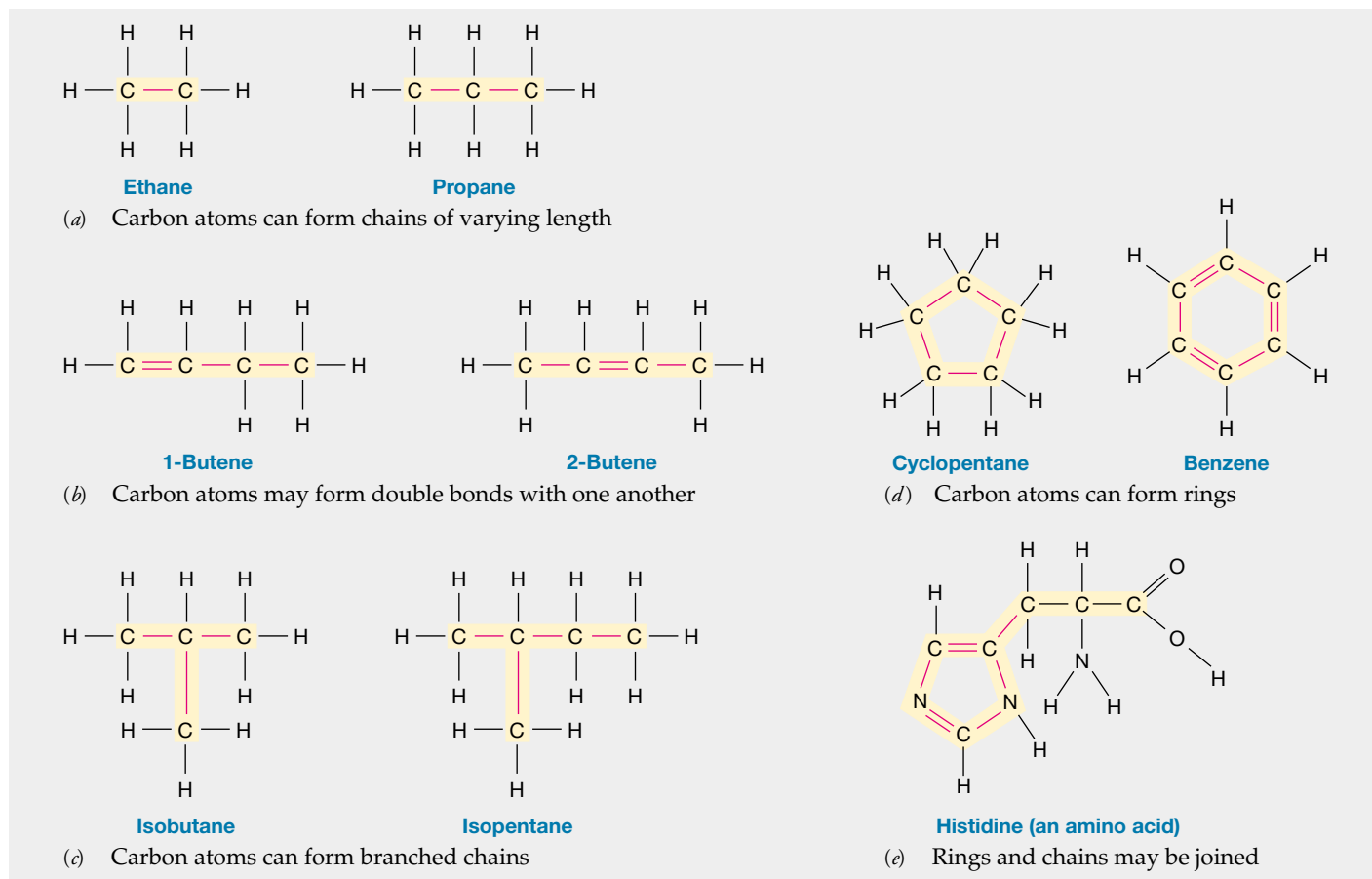


Figure 3-1 Organic molecules. Note that each carbon atom forms four covalent bonds, producing a wide variety of shapes.

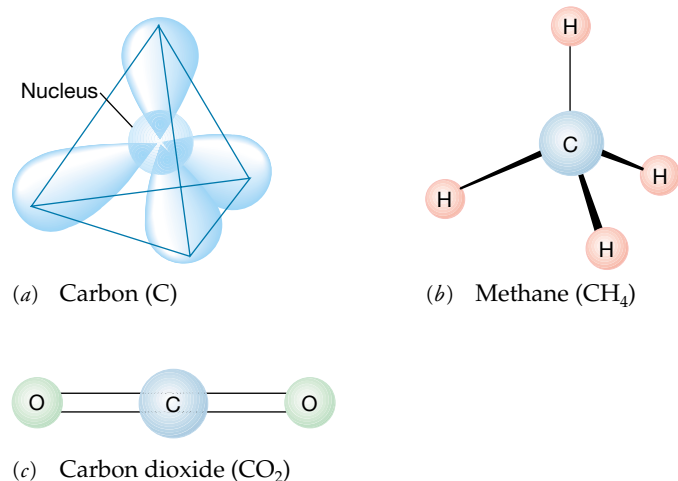
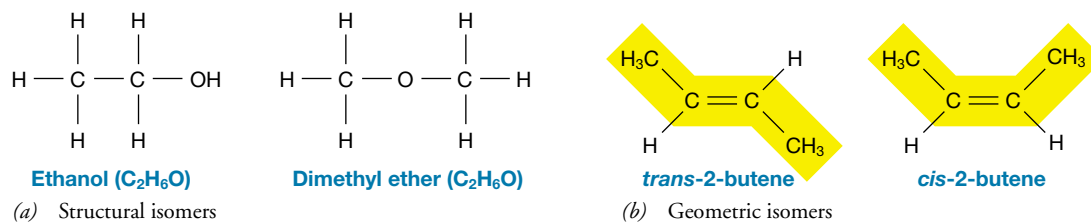


Figure 3-2 Carbon bonding. The three-dimensional arrangement of the bonds of a carbon atom (a) is responsible for the tetrahedral architecture of methane (b). (c) In carbon dioxide oxygen atoms are joined to a central carbon by polar double bonds. Because the bonds are arranged linearly, carbon dioxide does not have a positively charged end and a negatively charged end, and is therefore a nonpolar molecule.

Geometric isomers are compounds that are identical in the arrangement of their covalent bonds, but different in the spatial arrangement of groups of atoms. Geometric isomers are present in some compounds with carbon-to-carbon double bonds. Because double bonds are not flexible like single bonds, atoms joined to the carbons of a double bond cannot rotate freely about the axis of the bonds. The *cis-trans* isomers may be drawn as shown in Figure 3-3b. The designation *cis* (Latin, “on this side”) indicates that the two larger components are on the same side of the double bond. If they are on opposite sides of the double bond, the compound is designated a *trans* (Latin, “across”) isomer.

Enantiomers are molecules that are mirror images of one another (Fig. 3-3c). Recall that the four groups bonded to a single carbon atom are arranged at the vertices of a tetrahedron. If the four bonded groups are all different, the central carbon is described as asymmetric. Figure 3-3d illustrates that the four groups can be arranged about the asymmetric carbon in two different ways that are mirror images of each other. The two molecules are enantiomers if they cannot be superimposed on one another no matter how they are rotated in space. Although enantiomers have similar chemical properties and iden-



(c) Enantiomers are mirror images

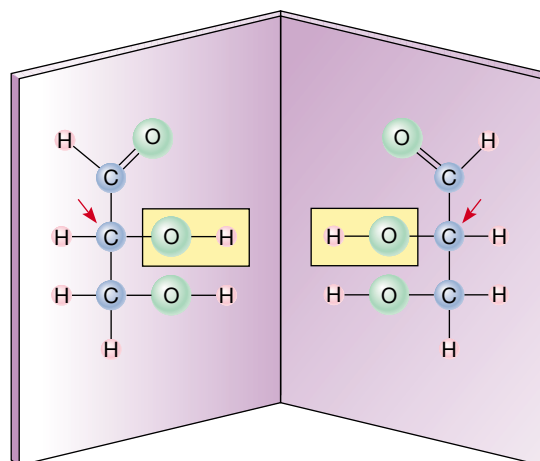


Figure 3–3 Isomers. Isomers have the same molecular formula, but their atoms are arranged differently. (a) Structural isomers differ in the covalent arrangement of their atoms. (b) Geometric, or *cis-trans*, isomers have identical covalent bonds but differ in the order in which groups of atoms are arranged in space. (c) and (d) Enantiomers are isomers that are mirror images of one another. The arrows indicate the asymmetric carbons. (c, Dennis Drenner)

tical physical properties, cells recognize the difference in shape and usually only one form is found in organisms.

The existence of isomers is not the only source of variety among organic molecules. The addition of various combinations of atoms, known as functional groups, can generate a vast array of molecules with differing properties.

FUNCTIONAL GROUPS CHANGE THE PROPERTIES OF ORGANIC MOLECULES

Because covalent bonds between hydrogen and carbon are non-polar, hydrocarbons lack distinct charged regions. For this reason, hydrocarbons are insoluble in water and tend to cluster together, through **hydrophobic** (“water-fearing”) **interactions**. This is not because bonds form among the hydrocarbons, but rather because the hydrogen-bonded water molecules exclude them and, in a sense, drive them together. Because of their hydrophobic nature, hydrocarbons do not interact readily with most other compounds.

However, the characteristics of an organic molecule can be changed dramatically by replacing one of the hydrogens with a group of atoms known as a **functional group**. Functional groups help determine the types of chemical reactions in which the compound participates. Most functional groups readily form associations, such as ionic and hydrogen bonds, with other molecules. Polar and ionic functional groups are hydrophilic because they associate strongly with polar water molecules.

The properties of the major classes of organic compounds—carbohydrates, lipids, proteins, and nucleic acids—are largely a consequence of the types and arrangement of functional groups they contain. When we know what kinds of functional groups are present in an organic compound, we can predict its chemical behavior. As you read the rest of this section please refer to Table 3–1 for the complete structural formulas of these groups, as well as additional information. Note that the symbol R is used to represent the *remainder* of the molecule of which each functional group is a part.

The **hydroxyl group** (abbreviated R—OH) must not be confused with the hydroxide ion (OH[−]) discussed in Chapter 2. The hydroxyl group is polar due to the presence of a

TABLE 3-1 Some Biologically Important Functional Groups

Functional group	Structural formula	Class of compounds characterized by group	Description
Hydroxyl	$R-OH$	Alcohols $\begin{array}{c} H & H \\ & \\ H-C & -C-OH \\ & \\ H & H \end{array}$ Ethanol	Polar because electronegative oxygen attracts covalent electrons
Carbonyl	$R-\overset{O}{\parallel}C-H$	Aldehydes $\begin{array}{c} O \\ \\ H-C-H \end{array}$ Formaldehyde	Carbonyl group carbon is bonded to at least one H atom; polar because electronegative oxygen attracts covalent electrons
	$R-\overset{O}{\parallel}C-R$	Ketones $\begin{array}{c} H & O & H \\ & & \\ H-C & -C & -C-H \\ & & \\ H & & H \end{array}$ Acetone	
Carboxyl	$R-\overset{O}{\parallel}C-OH$ Nonionized	Carboxylic acids (organic acids)	Weakly acidic; can release an H^+ ion
	$R-\overset{O}{\parallel}C-O^- + H^+$ Ionized	$\begin{array}{c} NH_2 & O \\ & \\ R-C & -C-OH \\ \\ H \end{array}$ Amino acid	
Amino	$R-N\begin{array}{l} /H \\ /H \end{array}$ Nonionized	Amines	Weakly basic; can accept an H^+ ion
	$R-N^+\begin{array}{l} /H \\ /H \\ /H \end{array}$ Ionized	$\begin{array}{c} NH_2 & O \\ & \\ R-C & -C-OH \\ \\ H \end{array}$ Amino acid	
Methyl	$R-CH_3$	Component of many organic compounds $\begin{array}{c} H & H \\ & \\ H-C & -C-H \\ & \\ H & H \end{array}$ Ethane	Hydrocarbon; nonpolar

TABLE 3-1 continued

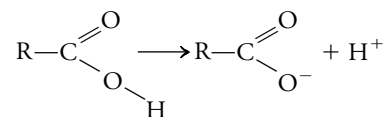
Functional group	Structural formula	Class of compounds characterized by group	Description
Phosphate	$\begin{array}{c} \text{O} \\ \parallel \\ \text{R}-\text{O}-\text{P}-\text{OH} \\ \\ \text{OH} \end{array}$ Nonionized	Organic phosphates	Weakly acidic; one or two H ⁺ ions can be released
	$\begin{array}{c} \text{O} \\ \parallel \\ \text{HO}-\text{P}-\text{O}-\text{R} \\ \\ \text{OH} \end{array}$ Phosphate ester (as found in ATP)		
	$\begin{array}{c} \text{O} \\ \parallel \\ \text{R}-\text{O}-\text{P}-\text{O}^- \\ \\ \text{O}^- \end{array} + 2 \text{H}^+$ Ionized		
Sulfhydryl	R—SH	Thiols $\begin{array}{c} \text{H} \quad \text{H} \quad \text{O} \\ \quad \quad \parallel \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{OH} \\ \quad \\ \text{SH} \quad \text{NH}_2 \end{array}$ Cysteine	Helps stabilize internal structure of proteins

strongly electronegative oxygen atom. If a hydroxyl group replaces one of the hydrogens of a hydrocarbon, the resulting molecule can have significantly altered properties. For example, ethane (Fig. 3-1*a*) is a hydrocarbon that is a gas at room temperature. If a hydrogen is replaced by a hydroxyl group, the resulting molecule is ethyl alcohol, or ethanol, which is found in alcoholic beverages (Fig. 3-3*a*). Ethanol is somewhat cohesive because the polar hydroxyl groups of adjacent molecules interact; it is therefore liquid at room temperature. Unlike ethane, ethyl alcohol can dissolve in water because the polar hydroxyl groups interact with the polar water molecules.

The **carbonyl group** consists of a carbon atom that has a double covalent bond with an oxygen atom. This double bond is polar because of the electronegativity of the oxygen; thus the carbonyl group is hydrophilic. The position of the carbonyl group in the molecule determines the class to which the molecule belongs. An **aldehyde** has a carbonyl group positioned at the end of the carbon skeleton (abbreviated R—CHO); a **ketone** has an internal carbonyl group (abbreviated R—CO—R).

The **carboxyl group** (abbreviated R—COOH) consists of a carbon atom joined by a double covalent bond to an oxygen atom, and by a single covalent bond to another oxygen, which is in turn bonded to a hydrogen atom. Two electronegative oxygen atoms in such close proximity establish an extremely polarized condition, which can cause the hydrogen

atom to be stripped of its electron and released as a hydrogen ion (H⁺). The carboxyl group then has one unit of negative charge (R—COO⁻):



Carboxyl groups are weakly acidic; only a fraction of the molecules ionize in this way. This group can therefore exist in one of two hydrophilic states: ionic or polar. Carboxyl groups are essential constituents of amino acids.

An **amino group** (abbreviated R—NH₂) includes a nitrogen atom covalently bonded to two hydrogen atoms. Amino groups are weakly basic because they are able to accept a hydrogen ion (proton), thus acquiring a unit of positive charge. Amino groups are components of amino acids and of nucleic acids.

A **phosphate group** (abbreviated R—PO₄H₂) is weakly acidic. The attraction of electrons by the oxygens can result in the release of one or two hydrogen ions, producing ionized forms with one or two units of negative charge. Phosphates are constituents of nucleic acids and certain lipids.

The **sulfhydryl group** (abbreviated R—SH), consisting of an atom of sulfur covalently bonded to a hydrogen atom,

is found in molecules called *thiols*. As we will see, amino acids that contain a sulfhydryl group can make important contributions to the structure of proteins.

The **methyl group** (abbreviated R—CH₃), a common hydrocarbon group, is nonpolar.

MANY BIOLOGICAL MOLECULES ARE POLYMERS

Many biological molecules such as proteins and nucleic acids are very large, consisting of thousands of atoms. Such giant molecules are known as **macromolecules**. Most macromolecules are **polymers**, produced by linking small organic compounds called **monomers** (Fig. 3–4). Just as all the words in this book have been written by arranging the 26 letters of the alphabet in various combinations, monomers can be grouped together to form an almost infinite variety of larger molecules. Just as we use different words to convey information, cells use different molecules to convey information. The thousands of different complex organic compounds present in organisms are constructed from about 40 small, simple monomers. For example, the 20 monomers called amino acids can be linked end-to-end in countless ways to form the polymers we know as proteins.

Each organism is unique because of differences in the monomer sequence within its DNA, the polymer that constitutes the information in the genes. Cells and tissues within the same organism are also different, due to variations in their component polymers. Muscle tissue and brain tissue differ in large part because of differences in the types and sequences of amino acids in proteins. Ultimately this protein structure is dictated by the sequence of monomers within the DNA of the organism, which is expressed somewhat differently in each cell type.

Polymers can be degraded to their component monomers by **hydrolysis** (which means “to break with water”). In a reaction regulated by a specific enzyme,¹ a hydrogen from a water molecule attaches to one monomer, and a hydroxyl from water attaches to the adjacent monomer.

The synthetic process by which monomers are covalently linked is called **condensation**. Because the *equivalent* of a molecule of water is removed during the reactions that combine monomers, the term *dehydration synthesis* is sometimes used to describe the process. However, in biological systems the synthesis of a polymer is not simply the reverse of hydrolysis, even though the net effect is the opposite of hydrolysis. Synthetic processes require energy and are regulated by different enzymes.

In the following sections we will examine carbohydrates, lipids, proteins, and nucleic acids. Our discussion will begin

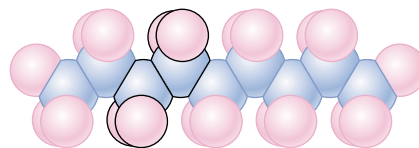


Figure 3–4 A simple polymer. This small polymer of polyethylene is formed by linking two-carbon ethylene (C₂H₄) monomers. One such monomer is outlined.

with the smaller, simpler forms of these compounds and will extend to the linking of these monomers to form macromolecules.

CARBOHYDRATES INCLUDE SUGARS, STARCHES, AND CELLULOSE

Sugars, starches, and cellulose are **carbohydrates**. Sugars and **starches** serve as energy sources for cells; **cellulose** is the main structural component of the walls that surround plant cells. Carbohydrates contain carbon, hydrogen, and oxygen atoms in a ratio of approximately one carbon to two hydrogens to one oxygen (CH₂O)_n. The term *carbohydrate*, meaning “hydrate (water) of carbon,” reflects the 2:1 ratio of hydrogen to oxygen, the same ratio found in water (H₂O). Carbohydrates contain one sugar unit (*monosaccharides*), two sugar units (*disaccharides*), or many sugar units (*polysaccharides*).

Monosaccharides are simple sugars

Monosaccharides typically contain from three to seven carbon atoms. In a monosaccharide, a hydroxyl group is bonded to each carbon except one; that carbon is double-bonded to an oxygen atom, forming a carbonyl group. If the carbonyl group is at the end of the chain, the monosaccharide is an aldehyde; if the carbonyl group is at any other position, the monosaccharide is a ketone. (By convention, the numbering of the carbon skeleton of a sugar begins with the carbon at or nearest the carbonyl end of the open chain.) The large number of polar hydroxyl groups, plus the carbonyl group, gives a monosaccharide hydrophilic properties.

Figure 3–5 shows simplified, two-dimensional representations of some common monosaccharides. The simplest carbohydrates are the three-carbon sugars (trioses): glyceraldehyde and dihydroxyacetone. Ribose and deoxyribose are common pentoses, sugars that contain five carbons; they are components of nucleic acids (DNA, RNA, and related compounds). Glucose, fructose, galactose, and other six-carbon sugars are called **hexoses**. (Note that the names of carbohydrates typically end in *-ose*.)

Glucose (C₆H₁₂O₆), the most abundant monosaccharide, is used as an energy source in most organisms. During cellular

¹An enzyme (see Chapter 6) is a protein catalyst that accelerates a specific chemical reaction.

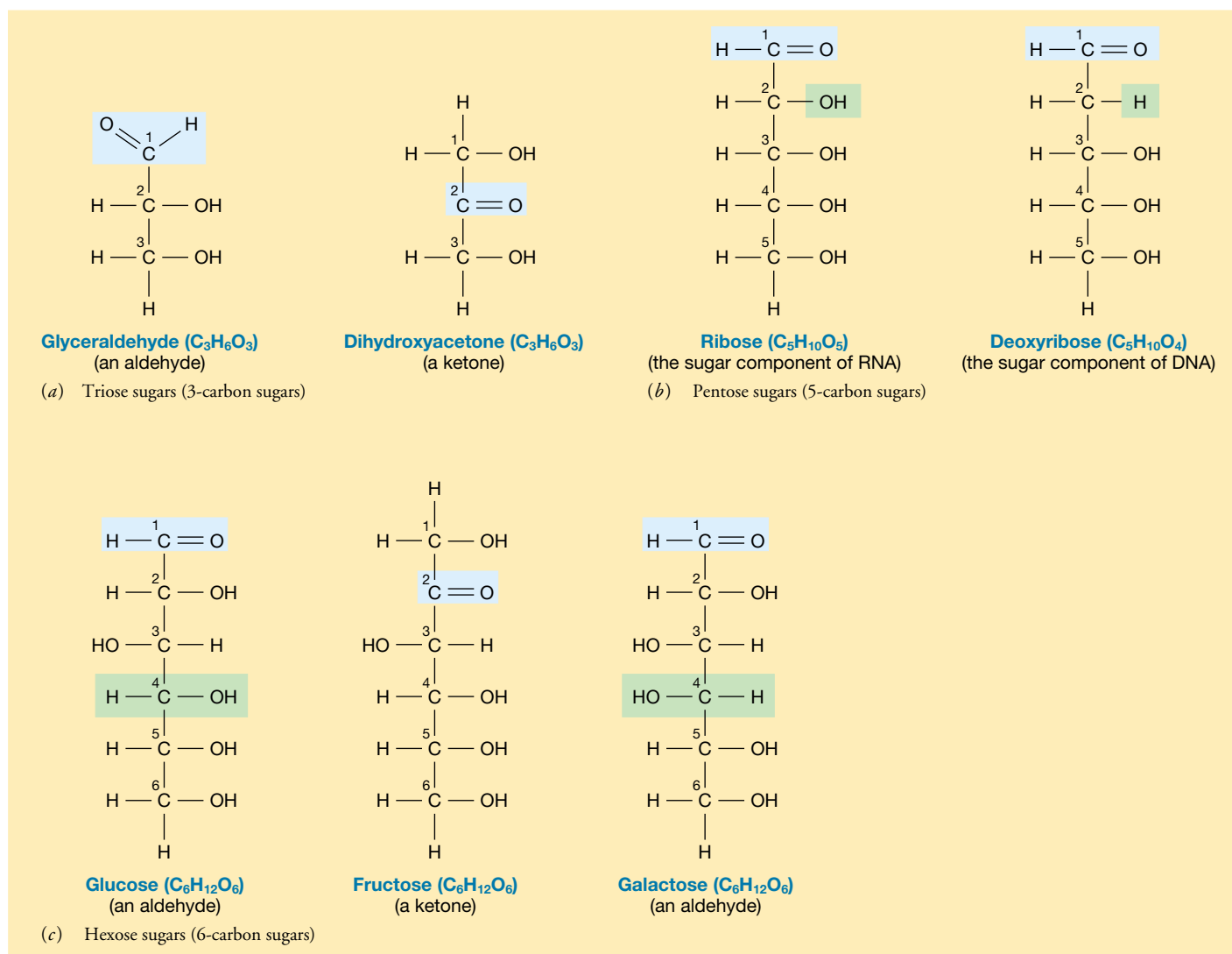


Figure 3-5 Monosaccharides. Shown are two-dimensional chain structures of (a) 3-carbon trioses, (b) 5-carbon pentoses, and (c) 6-carbon hexoses. Although it is convenient to show monosaccharides in this form, the pentoses and hexoses are more accurately depicted as ring structures, as in Figure 3-6.

respiration (Chapter 7), cells oxidize glucose molecules, converting the stored energy to a form that can be readily used for cellular work. Glucose is also used as a component in the synthesis of other types of compounds such as amino acids and fatty acids. So important is glucose in metabolism that its concentration is carefully kept at a homeostatic (relatively constant) level in the blood of humans and other complex animals.

Glucose and fructose are structural isomers: they have identical molecular formulas, but their atoms are arranged differently. In fructose (a ketone) the double-bonded oxygen is linked to a carbon within the chain rather than to a terminal carbon as in glucose (which is an aldehyde). Because of these differences, the two sugars have different properties. For example, fructose tastes sweeter than glucose.

Glucose and galactose are both hexoses and aldehydes.

However, they are mirror images (enantiomers) because they differ in the arrangement of the atoms attached to asymmetric carbon atom 4.

The “stick” formulas in Figure 3-5 give a clear but somewhat unrealistic picture of the structures of some common monosaccharides. As has been discussed, molecules are not two-dimensional; in fact, the properties of each compound depend largely on its three-dimensional structure. Thus, three-dimensional formulas are helpful in understanding the relationship between molecular structure and biological function. Molecules of glucose and other pentoses and hexoses in solution are actually rings, rather than extended straight carbon chains.

Glucose in solution (as in the cell) typically exists as a ring of five carbons and one oxygen. It assumes this configuration

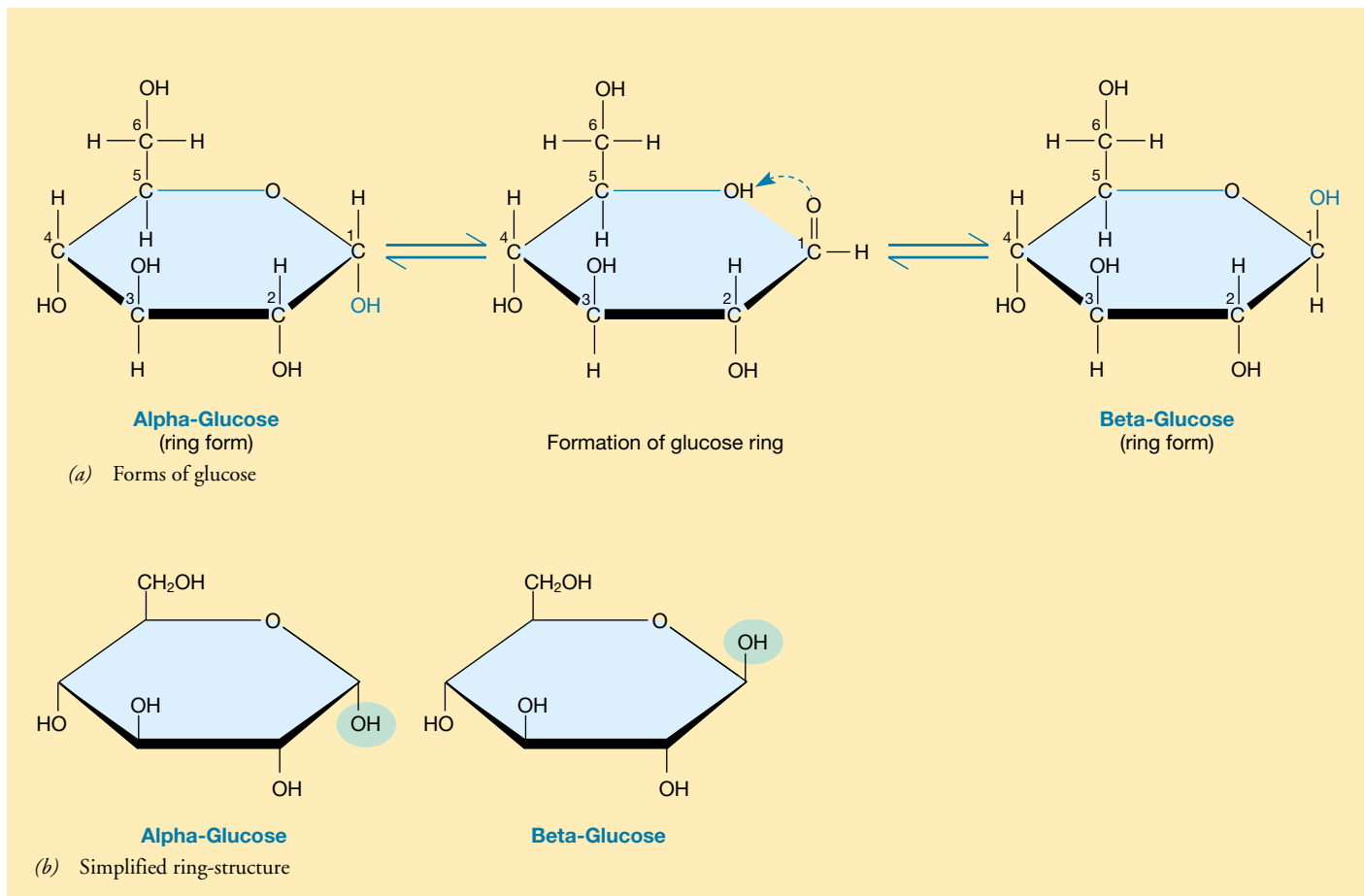


Figure 3–6 α and β forms of glucose. (a) When dissolved in water, glucose undergoes a rearrangement of its atoms, forming one of two possible ring structures, α -glucose or β -glucose. Although the drawing does not attempt to show the complete three-dimensional structure, the thick, tapered bonds in the lower portion of each ring represent the part of the molecule that would project out of the page toward you. (b) The essential differences between α -glucose and β -glucose are more readily apparent in these simplified structures. By convention, a carbon atom is assumed to be present at each angle in the ring unless another atom is shown. Most hydrogen atoms have been omitted.

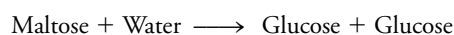
when its atoms undergo a rearrangement, permitting a covalent bond to connect carbon 1 to the oxygen attached to carbon 5 (Fig. 3–6). When glucose forms a ring, two isomeric forms are possible, differing only in the orientation of the hydroxyl ($-\text{OH}$) group attached to carbon 1. When this hydroxyl group is on the same side of the plane of the ring as the $-\text{CH}_2\text{OH}$ side group, the glucose is designated β -glucose. When it is on the side (with respect to the plane of the ring) opposite the $-\text{CH}_2\text{OH}$ side group, the compound is designated α -glucose.

Disaccharides consist of two monosaccharide units

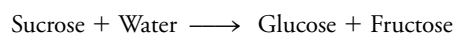
A disaccharide (two sugars) contains two monosaccharide rings joined by a **glycosidic linkage**, consisting of a central oxygen

covalently bonded to two carbons, one in each ring (Figure 3–7). The glycosidic linkage of a disaccharide generally forms between carbon 1 of one molecule and carbon 4 of the other molecule. The disaccharide maltose (malt sugar) consists of two covalently linked α -glucose units. Sucrose, common table sugar, consists of a glucose unit combined with a fructose unit. Lactose (the sugar present in milk) is composed of one molecule of glucose and one of galactose.

A disaccharide can be hydrolyzed, that is, split by the addition of water, into two monosaccharide units. During digestion, maltose is hydrolyzed to form two molecules of glucose:



Similarly, sucrose is hydrolyzed to form glucose and fructose:



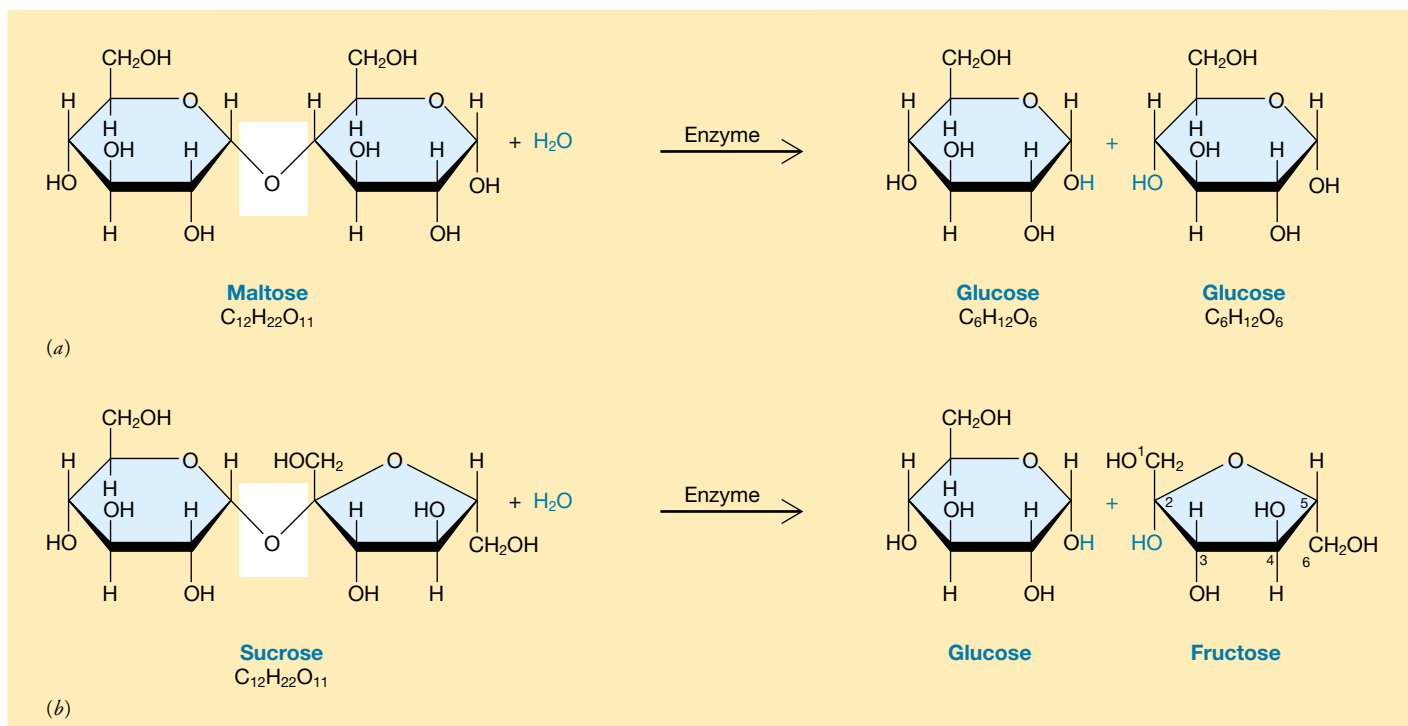


Figure 3-7 Hydrolysis of disaccharides. (a) Maltose may be broken down (as it is during digestion) to form two molecules of glucose. The highlighted glycosidic linkage is broken in a hydrolysis reaction, which requires the addition of water. (b) Sucrose can be hydrolyzed to yield a molecule of glucose and a molecule of fructose. Note that an enzyme (a protein catalyst) is needed to promote these reactions.

Polysaccharides can store energy or provide structure

The most abundant carbohydrates are the **polysaccharides**, a group that includes starches, glycogen, and cellulose. A polysaccharide is a macromolecule consisting of repeating units of simple sugars, usually glucose. Although the precise number of sugar units varies, thousands of units are typically present in a single molecule. The polysaccharide may be a single long chain or a branched chain. Because they are composed of different isomers and because the units may be arranged differently, polysaccharides vary in their properties. Those that can be easily broken down to their subunits are well suited for energy storage, whereas the macromolecular three-dimensional architecture of others makes them particularly well suited to form stable structures.

Starch is the main storage carbohydrate of plants

Starch, the typical form of carbohydrate used for energy storage in plants, is a polymer consisting of α -glucose subunits. These monomers are joined by α 1—4 linkages, which means that carbon 1 of one glucose is linked to carbon 4 of the next glucose in the chain (Fig. 3-8). Starch occurs in two forms,

amylose and amylopectin. Amylose, the simpler form, is unbranched. Amylopectin, the more common form, usually consists of about 1000 units in a branched chain.

Plant cells store starch mainly as granules within specialized organelles called **amyloplasts** (Fig. 3-8a). When energy is needed for cellular work, the plant can hydrolyze the starch, releasing the glucose subunits. Humans and other animals that eat plant foods have enzymes to hydrolyze starch.

Glycogen is the main storage carbohydrate of animals

Glycogen (sometimes referred to as *animal starch*) is the form in which glucose is stored as an energy source in animal tissues. It is similar in structure to plant starch, but more extensively branched and more water-soluble. Glycogen is stored mainly in liver and muscle cells.

Cellulose is a structural carbohydrate

Carbohydrates are the most abundant group of organic compounds on Earth, and **cellulose** is the most abundant carbohydrate; it accounts for 50% or more of all the carbon in plants (Fig. 3-9). Cellulose is a structural carbohydrate. Wood is about half cellulose, and cotton is at least 90% cellulose. Plant

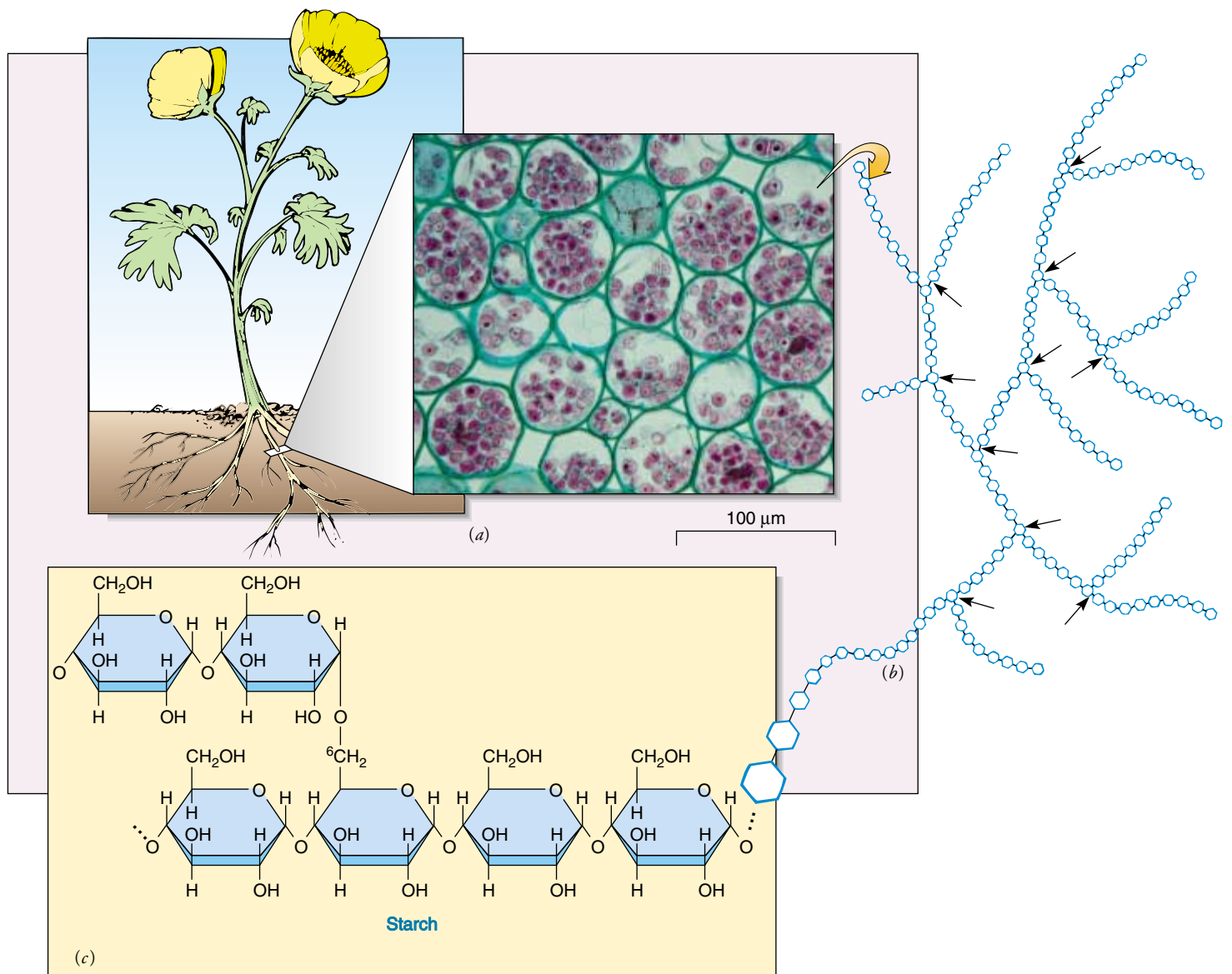


Figure 3-8 Starch, a storage polysaccharide. (a) Starch (stained purple) is stored in specialized organelles, called *amyloplasts*, in these cells of a buttercup root. (b) Starch is made up of highly branched chains; the arrows indicate the branch points. Each chain is actually in the form of a coil or helix, stabilized by hydrogen bonds between the glucose subunits. (c) Starch is composed of α -glucose molecules joined by glycosidic bonds. At the branch points are bonds between carbon 6 of the glucose in the straight chain and carbon 1 of the glucose in the branching chain. (a, Ed Reschke)

cells are surrounded by strong supporting cell walls consisting mainly of cellulose.

Cellulose is an insoluble polysaccharide composed of many glucose molecules joined together. The bonds joining these sugar units are different from those in starch. Recall that starch is composed of α -glucose subunits, joined by α 1—4 glycosidic linkages. Cellulose contains β -glucose monomers joined by β 1—4 linkages. These bonds are not split by the enzymes that hydrolyze the alpha linkages in starch. Humans, like most organisms, do not have enzymes that can digest cellulose and therefore cannot use it as a nutrient. Because cel-

lulose remains fibrous, as discussed in Chapter 45, it helps keep the digestive tract functioning properly.

Some microorganisms can digest cellulose to glucose. In fact, cellulose-digesting bacteria live in the digestive systems of cows and sheep, enabling these grass-eating animals to obtain nourishment from cellulose. Similarly, the digestive systems of termites contain microorganisms that digest cellulose (see Fig. 24-5a).

Cellulose molecules have characteristics that make them well suited for a structural role. The β -glucose subunits are joined in a way that allows extensive hydrogen bonding among

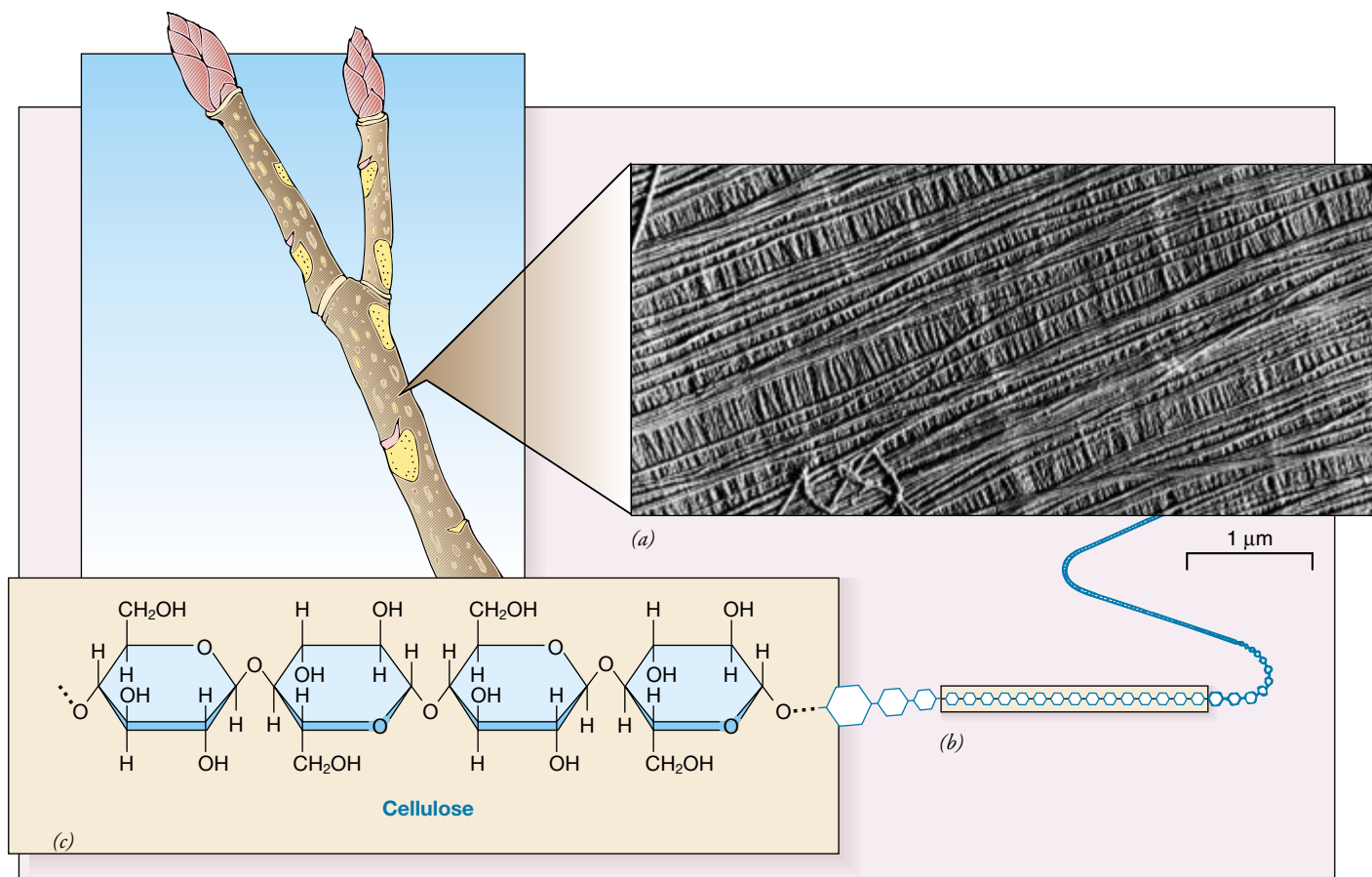


Figure 3–9 Cellulose, a structural polysaccharide. (a) An electron micrograph of cellulose fibers from a cell wall. The fibers visible in the photograph consist of bundles of cellulose molecules, interacting through hydrogen bonds. (b and c) The cellulose molecule is an unbranched polysaccharide composed of approximately 10,000 β -glucose units joined by glycosidic bonds. (a, Omikron/Photo Researchers, Inc.)

different cellulose molecules. Thus, cellulose molecules aggregate in long bundles of fibers, as shown in Figure 3–9a.

Some modified and complex carbohydrates have special roles

Many derivatives of monosaccharides are important biological molecules. Some form important structural components. The amino sugars galactosamine and glucosamine are compounds in which a hydroxyl group ($-\text{OH}$) is replaced by an amino group ($-\text{NH}_2$). Galactosamine is present in cartilage, a constituent of the skeletal system of vertebrates. *N*-acetyl glucosamine (NAG) subunits, joined by glycosidic bonds, compose **chitin**, a main component of the external skeletons of insects, crayfish, and other arthropods (Fig. 3–10), and of the cell walls of fungi. Chitin forms very tough structures because, as in cellulose, its molecules interact through multiple hydrogen bonds.

Carbohydrates may also be combined with proteins to form **glycoproteins**, compounds present on the outer surface

of cells other than bacteria. Some of these carbohydrate chains allow cells to adhere to one another, while others provide protection. Most proteins secreted by cells are glycoproteins. Carbohydrates can combine with lipids to form **glycolipids**, compounds present on the surfaces of animal cells that are thought to allow cells to recognize and interact with one another (see Chapters 4 and 5).

LIPIDS ARE FATS OR FATLIKE SUBSTANCES

Lipids are a heterogeneous group of compounds that are defined, not by their structure, but rather by the fact that they are soluble in nonpolar solvents (ether, chloroform, etc.) and are relatively insoluble in water. Lipid molecules have these properties because they consist mainly of carbon and hydrogen, with few oxygen-containing functional groups. Oxygen

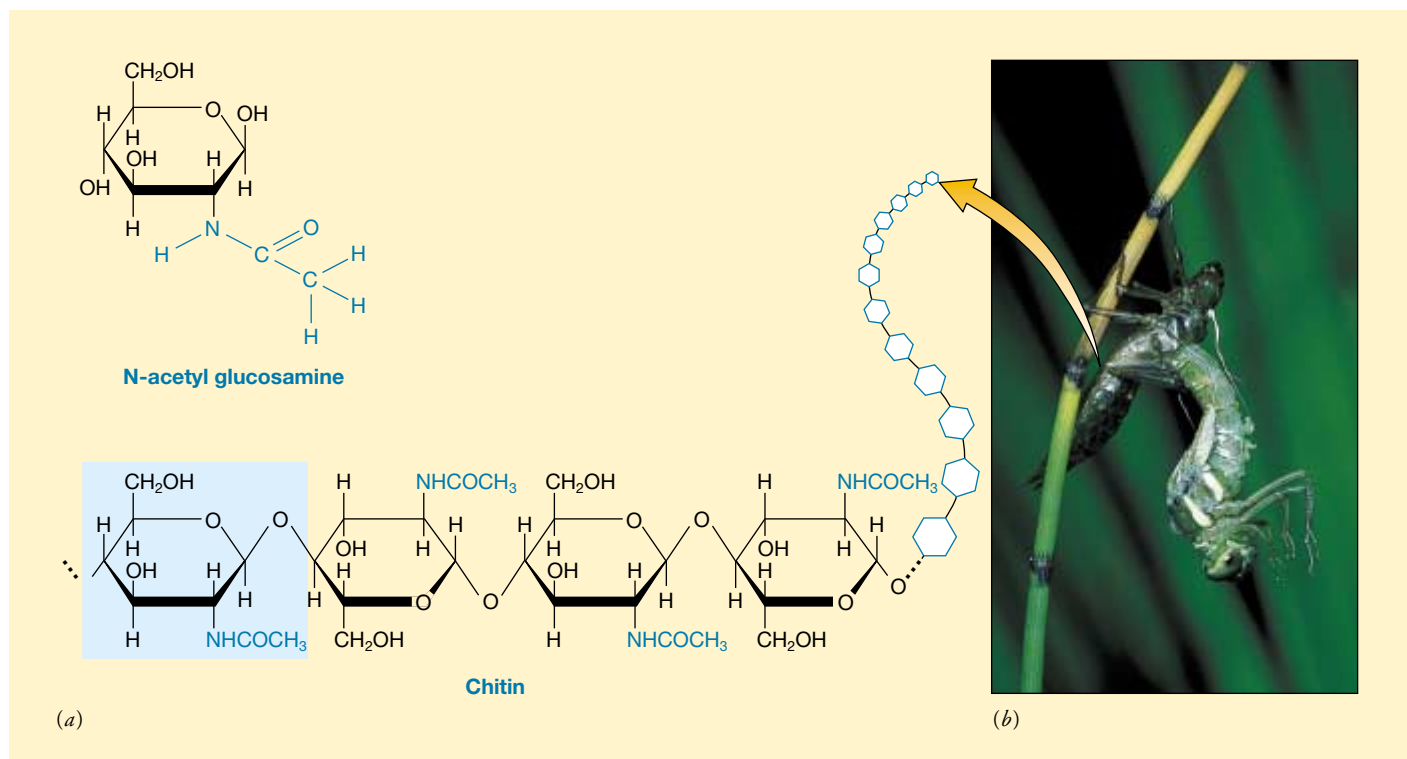


Figure 3–10 Chitin, a structural polysaccharide. (a) Chitin is a polymer composed of *N*-acetyl glucosamine (NAG) subunits. (b) Chitin is an important component of the exoskeleton (outer covering) that this dragonfly is shedding. (b, Dwight R. Kuhn)

atoms are commonly found in hydrophilic functional groups; therefore lipids, with little oxygen, tend to be hydrophobic. Among the biologically important groups of lipids are neutral fats, phospholipids, carotenoids (orange and yellow plant pigments), steroids, and waxes. Some lipids are used for energy storage, others serve as structural components of cellular membranes, and some are important hormones.

Neutral fats contain glycerol and fatty acids

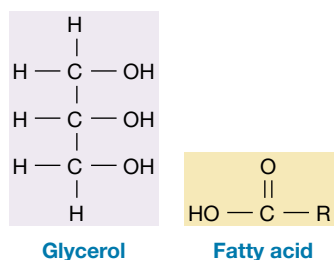
The most abundant lipids in living organisms are the **neutral fats**. These compounds are an economical form of reserve fuel storage because they yield more than twice as much energy per gram as do carbohydrates. Carbohydrates and proteins can be transformed by enzymes into fats and stored within the cells of adipose (fat) tissue of animals and in some seeds and fruits of plants.

A neutral fat consists of glycerol joined to one, two, or three fatty acids. **Glycerol** is a three-carbon alcohol that contains three hydroxyl (—OH) groups (Fig. 3–11). A **fatty acid** is a long, unbranched hydrocarbon chain with a carboxyl group (—COOH) at one end. About 30 different fatty acids are commonly found in lipids, and they typically have an even number of carbon atoms. For example, butyric acid, present in ran-

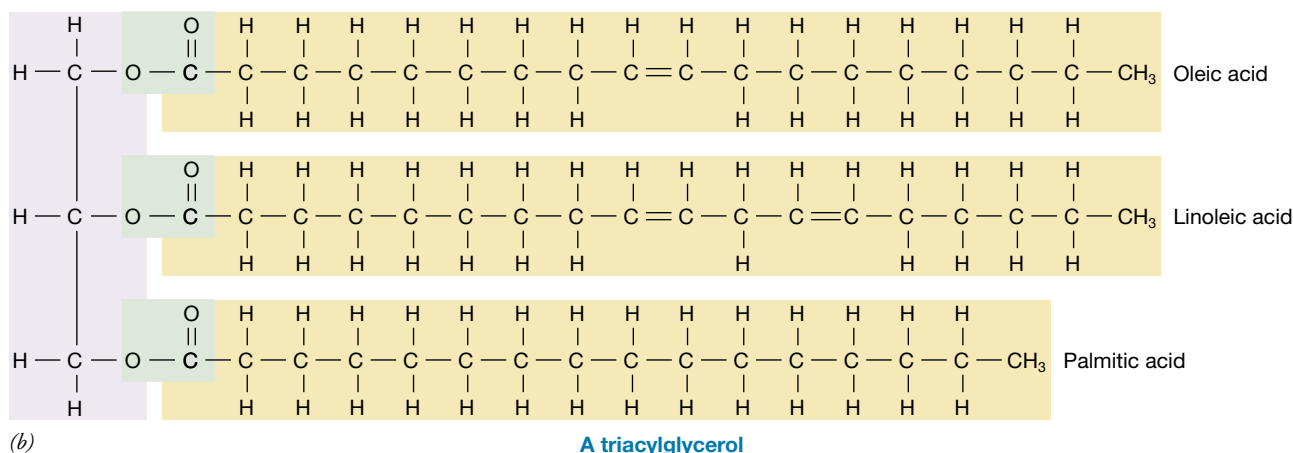
cid butter, has four carbon atoms. Oleic acid, with 18 carbons, is the most widely distributed fatty acid in nature and is found in most animal and plant fats.

Saturated fatty acids contain the maximum possible number of hydrogen atoms. Fats high in saturated fatty acids, such as animal fat and solid vegetable shortening, tend to be solid at room temperature. This is because even electrically neutral, nonpolar molecules can develop transient regions of weak positive charge and weak negative charge, as the constant motion of their electrons causes some regions to have a temporary excess of electrons, while others have a temporary electron deficit. These slight opposite charges result in attractions, known as **van der Waals forces**, between adjacent molecules. Although van der Waals interactions are individually weak, they can be strong when many occur among long hydrocarbon chains.

Unsaturated fatty acids include one or more adjacent pairs of carbon atoms joined by a double bond. Therefore they are not fully saturated with hydrogen. Fatty acids with one double bond are called **monounsaturated fatty acids**, while those with more than one double bond are **polyunsaturated fatty acids**. Fats containing a high proportion of monounsaturated or polyunsaturated fatty acids tend to be liquid at room temperature. This is because each double bond produces a bend in the hydrocarbon chain that prevents it from aligning closely



(a)



(c)

Figure 3-11 Neutral fats. (a) Glycerol and fatty acids are the components of neutral fats. Every fatty acid contains a carboxyl ($-\text{COOH}$) group, plus a long hydrocarbon (represented by “R”) that is specific for that particular fatty acid. (b) Glycerol is attached to fatty acids by ester linkages, shown in green. Note that oleic and linoleic acid are unsaturated fatty acids. They are drawn as straight chains, but the molecules are actually bent or kinked wherever a carbon-to-carbon double bond appears. (c) Commonly used cooking fats contain triacylglycerols. (c, Kenneth Knott/Fine Light Photography)

with an adjacent chain, thereby limiting van der Waals interactions.

At least two unsaturated fatty acids (linoleic acid and arachidonic acid) are essential nutrients that must be obtained from food because the human body cannot synthesize them. However, the amounts required are small, and deficiencies are rarely seen. There is no dietary requirement for saturated fatty acids.

When a glycerol molecule combines chemically with one fatty acid, a **monoacylglycerol** (sometimes called *monoglyceride*) is formed. **Diacylglycerols** (or *diglycerides*) and **triacylglycerols** (or *triglycerides*) contain two or three fatty acids, respectively. In each condensation reaction, the equivalent of a water molecule is removed as one of the glycerol’s hydroxyl groups reacts with the carboxyl group of a fatty acid. Such a reaction between a hydroxyl group and a carboxyl group results in the formation of a covalent linkage known as an **ester linkage** (Fig. 3-11b). During digestion, neutral fats are hydrolyzed to produce fatty acids and glycerol (see Chapter 45).

Phospholipids are components of cellular membranes

Phospholipids belong to a group of lipids, called **amphipathic lipids**, in which one end of each molecule is hydrophilic and the other end is hydrophobic. The two ends of a phospholipid differ both physically and chemically. A phospholipid consists of a glycerol molecule attached at one end to two fatty acids, and at the other end to a phosphate group linked to an organic compound such as choline. The organic compound usually contains nitrogen (Fig. 3-12). (Note that phosphorus and nitrogen are absent in the neutral fats.) The fatty acid portion of the molecule (containing the two hydrocarbon “tails”) is hydrophobic and not soluble in water. However, the portion composed of glycerol, phosphate, and the organic base (the “head” of the molecule) is ionized and readily water-soluble. The amphipathic properties of these lipid molecules make them uniquely suited to form cellular membranes, structures that will be discussed in Chapters 4 and 5.

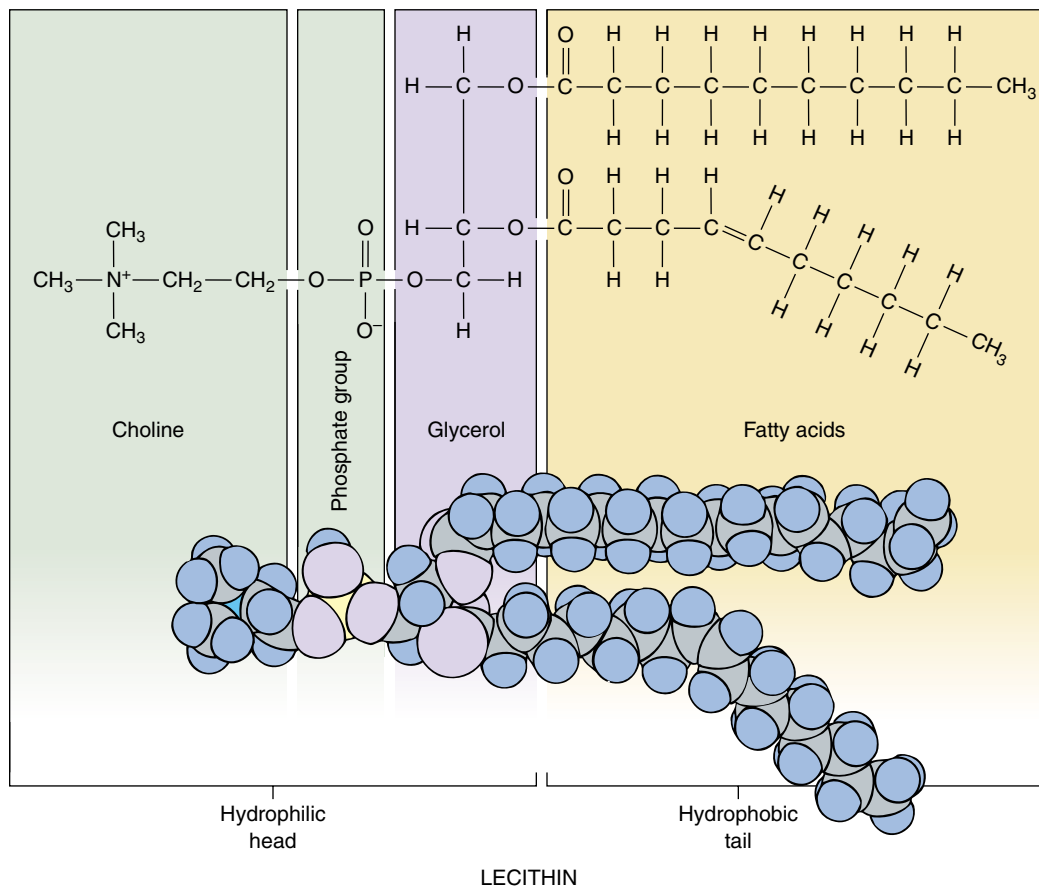


Figure 3–12 A phospholipid. Note the hydrophobic tail, made up of two fatty acids, and the hydrophilic head, which includes a glycerol bonded to a phosphate group, which is in turn bonded to an organic group that can vary. Choline is the organic group in the molecule shown, lecithin (or phosphatidylcholine). The lower fatty acid in the figure is monounsaturated; it contains one double bond that produces a characteristic bend in the chain.

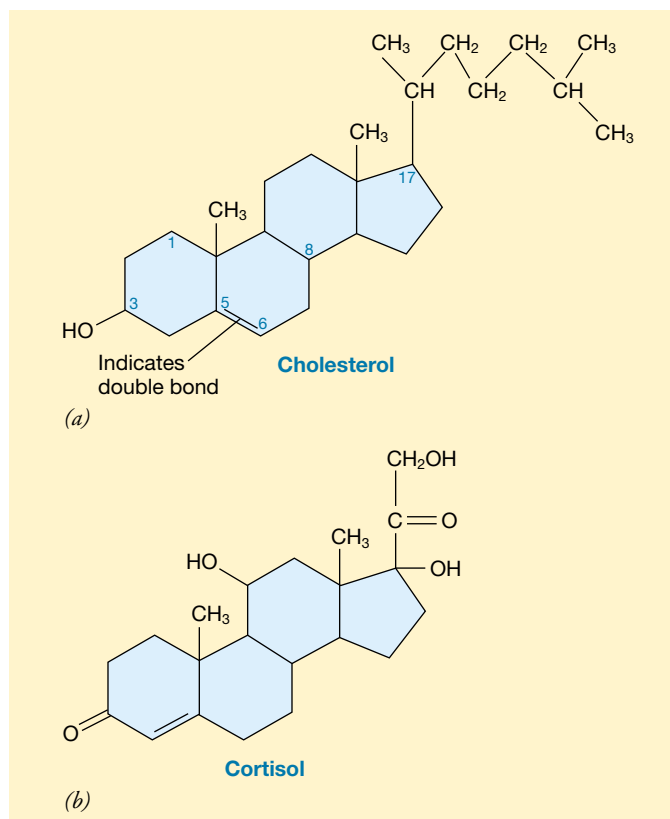
Carotenoid plant pigments are derived from isoprene units

The orange and yellow plant pigments called **carotenoids** are classified with the lipids because they are insoluble in water and have an oily consistency. These pigments, found in the cells of all plants, play a role in photosynthesis. As discussed in *Making the Connection: Molecules that Absorb Light*, carotenoid molecules and many other important pigments consist of five-carbon hydrocarbon monomers known as *isoprene units*.

Steroids contain four rings of carbon atoms

A **steroid** consists of carbon atoms arranged in four attached rings; three of the rings contain six carbon atoms, and the fourth contains five (Fig. 3–13). The length and structure of the side chains that extend from these rings distinguish one

Figure 3–13 Steroids. Four attached rings—three six-carbon rings and one with five carbons—make up a steroid. Note that some carbons are shared by two rings. In these simplified structures, a carbon atom is present at each angle of a ring; the hydrogen atoms attached directly to the ring have not been drawn. (a) Cholesterol is an essential component of animal cellular membranes. (b) Cortisol is a steroid hormone secreted by the adrenal glands. Notice that cortisol differs from cholesterol in its attached functional groups.



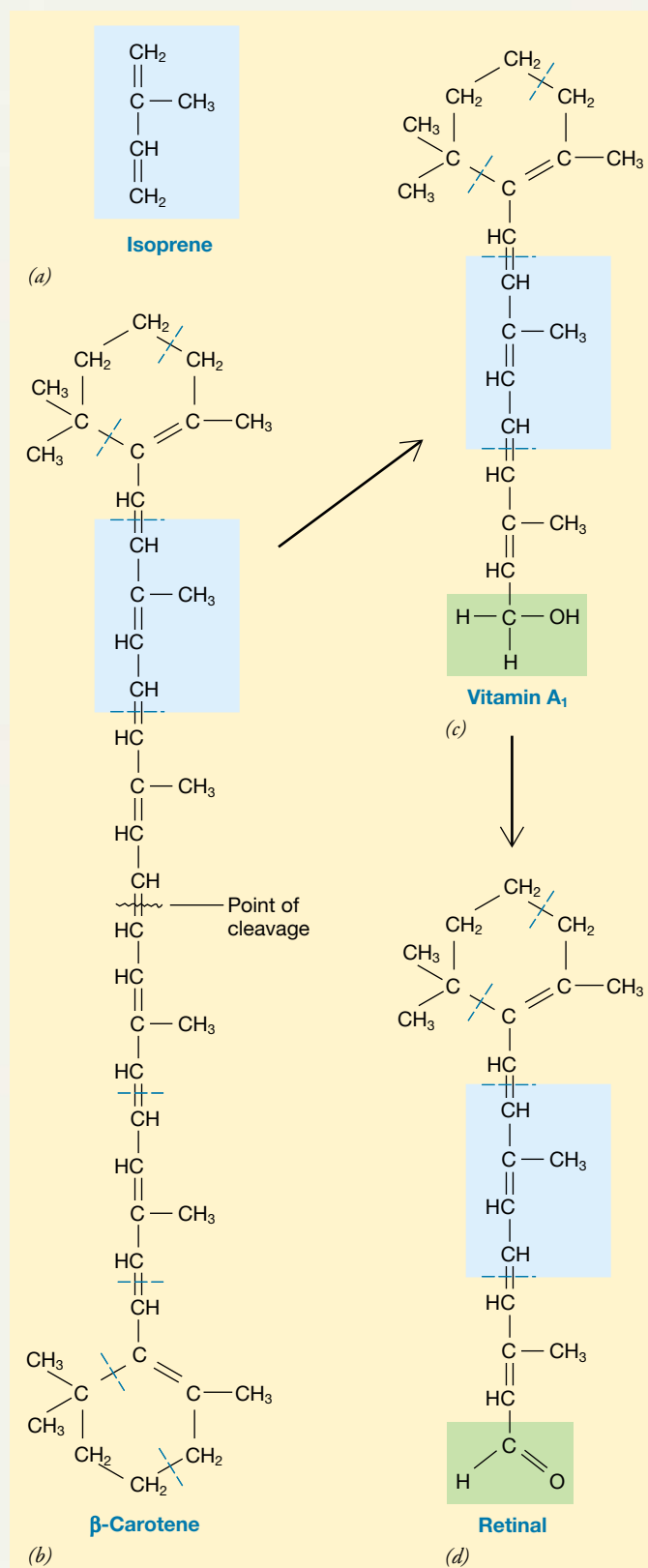
MAKING THE CONNECTION

MOLECULES THAT ABSORB LIGHT

What do some biologically important pigments have in common? Many are formed from 5-carbon isoprene monomers (*a*). One such molecule is β -carotene (*b*), a yellow plant pigment that plays a role in photosynthesis. (The dashed lines indicate the boundaries of the individual isoprene units within β -carotene.) Most animals can break carotenoids at the point of cleavage marked in the diagram and convert them to vitamin A (*c*). Vitamin A can then be converted to the visual pigment retinal (*d*). Retinal can exist in more than one, three-dimensional form (not shown).

Interestingly, eyes are found in three different lines of animals: the mollusks, insects, and vertebrates. There is no evidence that they share a common ancestor equipped with eyes, although they may have had a common ancestor with primitive photoreceptors. They all have the same compound involved in the process of light reception, retinal, a molecule that is apparently uniquely fitted for this role.

Notice that all these molecules have a pattern of double bonds alternating with single bonds. The electrons that make up these bonds can move about relatively easily when light strikes the molecule. Such molecules tend to be highly colored pigments because they strongly absorb light of certain wavelengths and reflect other wavelengths.



β carotene, vitamin A, and retinal are composed of isoprene subunits.

steroid from another. Steroids are synthesized from isoprene units.

Among the steroids of biological importance are cholesterol, bile salts, reproductive hormones, and cortisol and other hormones secreted by the adrenal cortex. Cholesterol is a structural component of animal cell membranes; plant cell membranes contain molecules similar to cholesterol. Bile salts emulsify fats in the intestine so that they can be enzymatically hydrolyzed. Steroid hormones regulate certain aspects of metabolism in a variety of animals, including vertebrates, insects, and crabs.

Some chemical mediators are derived from fatty acids

Animal cells secrete chemicals that permit them to communicate with each other or to regulate their own activities. Some chemical mediators are produced by the modification of fatty acids that have been removed from membrane phospholipids. These include prostaglandins, which have varied roles, including promoting inflammation and smooth muscle contraction. Certain hormones, such as the juvenile hormone of insects, are also fatty acid derivatives (Chapter 47).

PROTEINS ARE MACROMOLECULES FORMED FROM AMINO ACIDS

Because **proteins** are extraordinarily versatile macromolecules, they are of central importance in the chemistry of life. Proteins can be assembled into a variety of shapes, allowing them to serve as major structural components of cells and tissues. For this reason, growth and repair, as well as maintenance of the organism, depend on an adequate supply of these compounds. Most **enzymes** (molecules that speed up the thousands of different chemical reactions that take place in an organism) are proteins. Proteins serve in a great many other specialized capacities, including cellular motors, hormones and other chemical messengers, regulators of cellular activities, and defenders against foreign invaders.

The protein constituents of a cell are the clues to its lifestyle. Each cell type contains characteristic forms, distributions, and amounts of protein that largely determine what the cell looks like and how it functions. A muscle cell contains large amounts of the proteins myosin and actin, which are responsible for its appearance as well as for its ability to contract. The protein hemoglobin, found in red blood cells, is responsible for the specialized function of oxygen transport.

Although carbohydrates and lipids tend to have the same structures among different species, most proteins are species-specific; that is, their structures vary from species to species. The specific proteins present (determined by the instructions in the genes) are largely responsible for differences among species. Thus, the proteins in the cells of a dog vary somewhat from those of a fox and even more from those of an oak tree. The degree of difference in the proteins of two species is

thought to depend on evolutionary relationships. Distantly related organisms have proteins that differ more markedly than those of closely related forms.

Some proteins differ slightly even among individuals of the same species. Although it is common for individuals to have many identical proteins, most individuals are biochemically unique in the sense that no two are likely to have *all* identical proteins unless they are also genetically identical (i.e., identical twins or members of closely inbred strains).

Amino acids are the subunits of proteins

Amino acids, the constituents of proteins, have an amino group ($-\text{NH}_2$) and a carboxyl group ($-\text{COOH}$) bonded to the same asymmetric carbon atom, known as the **alpha carbon**. There are about 20 amino acids commonly found in proteins, each uniquely identified by the variable group (R group) bonded to the alpha carbon (Fig. 3–14). Glycine, the simplest amino acid, has a hydrogen atom as its R group; alanine has a methyl ($-\text{CH}_3$) group.

Amino acids in solution at neutral pH are mainly dipolar ions. This is generally how amino acids exist at cellular pH. Each carboxyl group ($-\text{COOH}$) donates a proton and becomes dissociated ($-\text{COO}^-$), while each amino group ($-\text{NH}_2$) accepts a proton and becomes $-\text{NH}_3^+$ (Fig. 3–15). Because of their amino and carboxyl groups, amino acids in solution resist changes in acidity and alkalinity and so are important biological buffers.

The amino acids are grouped in Figure 3–14 by the properties of their side chains. These broad groupings actually include amino acids with a fairly wide range of properties. Amino acids classified as having *nonpolar* side chains tend to have hydrophobic properties, whereas those classified as *polar* are more hydrophilic. An acidic amino acid has a side chain that contains a carboxyl group. At cellular pH the carboxyl group is dissociated, giving the R group a negative charge. A basic amino acid becomes positively charged when the amino group in its side chain accepts a hydrogen ion. Acidic and basic side chains are ionic at cellular pH and therefore hydrophilic.

In addition to the 20 common amino acids, some proteins have unusual ones. These rare amino acids are produced by the modification of common ones after they have become part of a protein. For example, lysine and proline may be converted to hydroxylysine and hydroxyproline after they have been incorporated into collagen. These amino acids can form cross links between the peptide chains that make up collagen. Such cross links are responsible for the firmness and great strength of the collagen molecule, which is a major component of cartilage, bone, and other connective tissues.

With some exceptions, bacteria and plants can synthesize all their needed amino acids from simpler substances. If the proper raw materials are available, the cells of humans and animals can manufacture some, but not all, of the biologically significant amino acids. Those that animals cannot synthesize and so must obtain from the diet are known as **essential amino acids**. Animals differ in their biosynthetic capacities; what is

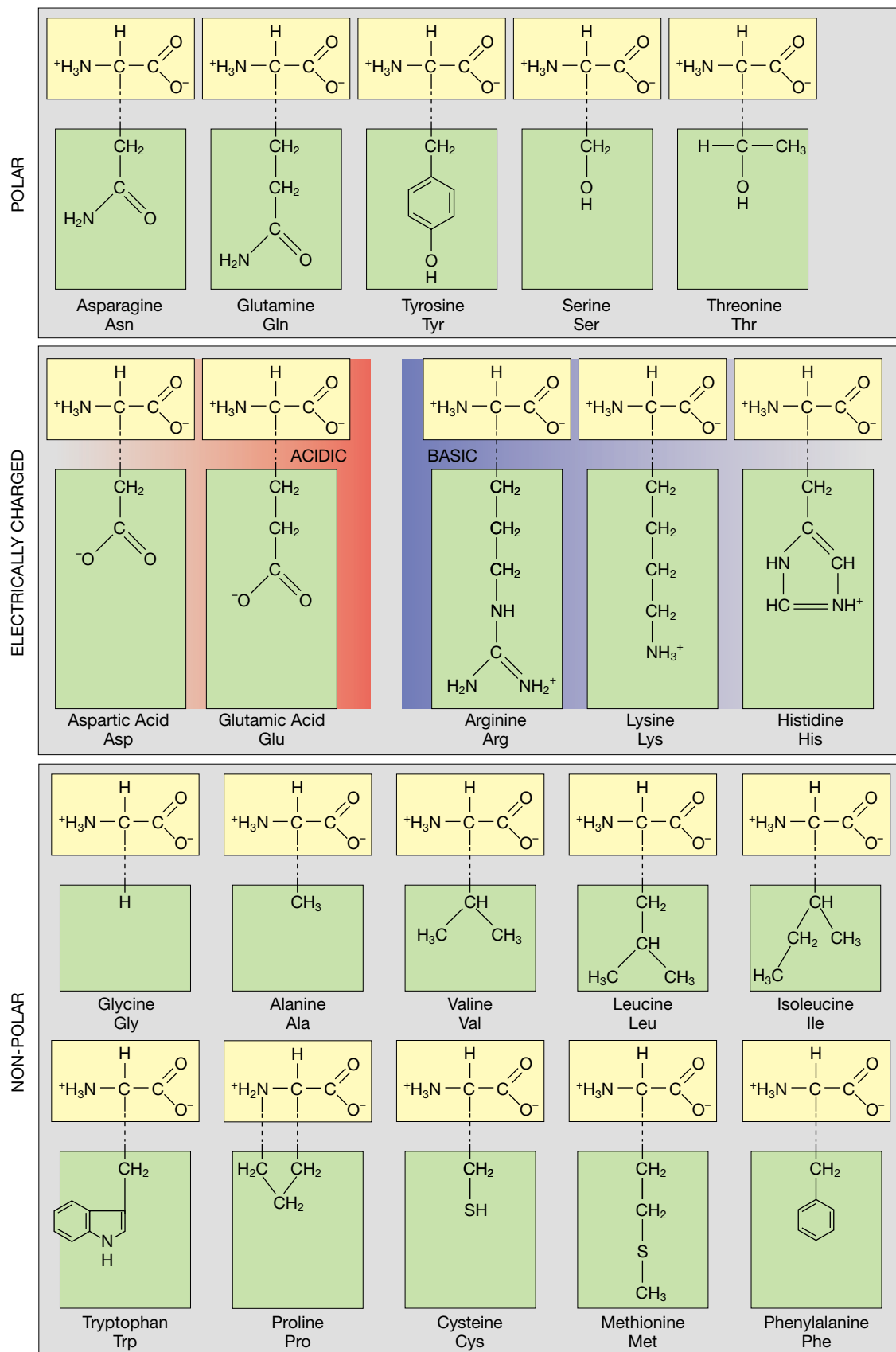


Figure 3-14 The twenty common amino acids. Amino acids designated *polar* have relatively hydrophilic side chains, while the side chains of those referred to as *nonpolar* are relatively hydrophobic. Carboxyl groups and amino groups are electrically charged at cellular pH; therefore acidic and basic amino acids are hydrophilic. The three-letter symbols are the conventional abbreviations for the amino acids.

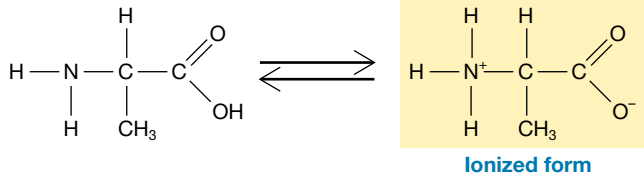


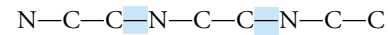
Figure 3–15 An amino acid at pH 7. In living cells, amino acids exist mainly as dipolar ions.

an essential amino acid for one species may not be for another. The essential amino acids for humans include isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, valine, histidine, and (in children) arginine.

Peptide bonds join amino acids

Amino acids combine chemically with one another by a condensation reaction that bonds the carboxyl carbon of one molecule to the amino nitrogen of another (Fig. 3–16). The covalent carbon-to-nitrogen bond linking two amino acids together is called a **peptide bond**. When two amino acids combine, a **dipeptide** is formed; a longer chain of amino acids is a **polypeptide**. A protein consists of one or more polypeptide chains. Each polypeptide has a free amino group at one end and a free carboxyl group (belonging to the last amino acid added to the chain) at the opposite end. The other amino and carboxyl groups of the amino acid monomers (except those in side chains) are part of the peptide bonds. The complex process by which polypeptides are synthesized is discussed in Chapter 12.

A polypeptide may contain hundreds of amino acids joined in a specific linear order. The backbone of the chain includes the repeating sequence



plus all other atoms *except those in the R groups*. The R groups of the amino acids extend from this backbone.

An almost infinite variety of protein molecules is possible, differing from one another in the number, types, and sequences of amino acids they contain. The 20 types of amino acids found in proteins may be thought of as letters of a protein alphabet; each protein is a very long sentence made up of amino acid letters.

Proteins have four levels of organization

The polypeptide chains making up a protein are twisted or folded to form a macromolecule with a specific *conformation*, or three-dimensional shape. Some polypeptide chains form long fibers. **Globular** proteins are tightly folded into compact, roughly spherical shapes. There is a close relationship between a protein's conformation and its function. For example, a typical enzyme is a globular protein with a unique shape that permits it to catalyze a specific chemical reaction. Similarly, the shape of a protein hormone enables it to combine with receptors on its target cell (the cell the hormone acts upon).

Four main levels of protein organization can be recognized: primary, secondary, tertiary, and quaternary. An analogy for secondary and tertiary structure is depicted in Figure 3–17.

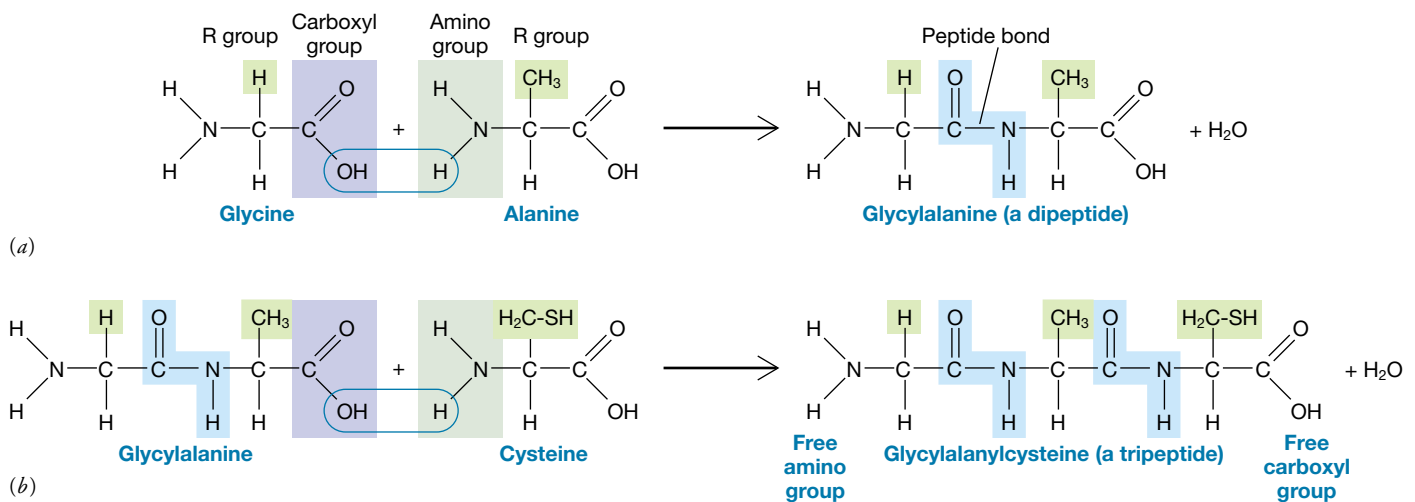


Figure 3–16 Peptide bonds. (a) A dipeptide is formed by the removal of the equivalent of a water molecule from the carboxyl group of one amino acid and the amino group of another amino acid. The resulting peptide bond is a covalent, carbon-to-nitrogen bond. (b) The carboxyl group of the dipeptide reacts with the amino group of a third amino acid to form a chain of three amino acids (a tripeptide, or small polypeptide). Additional amino acids can be added to form a long polypeptide chain with a free amino group at one end and a free carboxyl group at the other.

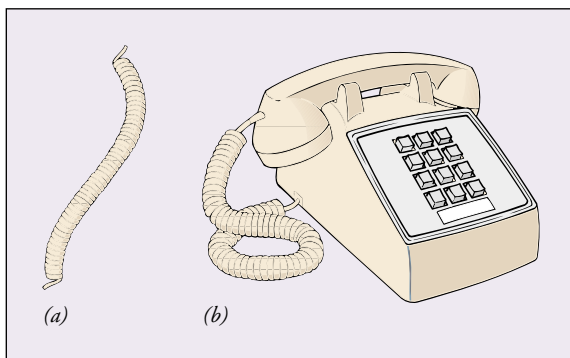


Figure 3-17 Levels of protein structure. A telephone cord provides a familiar analogy for the levels of protein structure. (a) Secondary structure, in this case analogous to an α -helix. (b) Tertiary structure, in which interactions among side chains cause the molecule to fold back on itself.

Primary structure is the amino acid sequence

The sequence of amino acids, joined by peptide bonds, is the **primary structure** of a polypeptide chain. As discussed in Chapter 12, this sequence is specified by the instructions in a gene. Using analytical methods developed in the early 1950s, investigators can determine the exact sequence of amino acids in a protein molecule. Insulin, a hormone secreted by the pancreas and used in the treatment of diabetes, was the first protein for which the exact sequence of amino acids in the polypeptide chains was identified. Insulin is a very small protein, consisting of 51 amino acid units in two linked chains, each with its own primary structure (Fig. 3-18).

Primary structure is always represented in a simple, linear, “beads-on-a-string” form. However, the overall conformation of a protein is far more complex, involving interactions among the various amino acids that comprise the primary structure of the molecule. Therefore, the higher orders of structure—secondary, tertiary, and quaternary—ultimately derive from the specific amino acid sequence (i.e., the primary structure).

Secondary structure results from hydrogen bonding involving the backbone

Some regions of a polypeptide exhibit **secondary structure**, which is highly regular because it is maintained by hydrogen

bonds between certain atoms of the polypeptide chain’s uniform backbone.

A common secondary structure in protein molecules is the **α -helix**, a region where a polypeptide chain forms a uniform spiral coil (Fig. 3-19a). The helical structure is determined and maintained by the formation of hydrogen bonds between the backbones of the amino acids in successive turns of the spiral coil. Each hydrogen bond forms between an oxygen with a partial negative charge and a hydrogen with a partial positive charge. The oxygen is part of the remnant of the carboxyl group of one amino acid; the hydrogen is part of the remnant of the amino group of the fourth amino acid down the chain. Thus 3.6 amino acids are included in each complete turn of the helix. Every amino acid in an α -helix is hydrogen-bonded in this way.

The α -helix is the basic structural unit of some fibrous proteins that make up wool, hair, skin, and nails. The elasticity of these fibers is due to a combination of physical factors (the helical shape) and chemical factors (hydrogen bonding). Although hydrogen bonds maintain the helical structure, these bonds can be broken, allowing the fibers to stretch under tension (like a telephone cord). When the tension is released, the fibers recoil and hydrogen bonds reform. This is why human hairs can stretch to some extent and then snap back to their original length.

Another type of secondary structure is the **β -pleated sheet**² (Fig. 3-19b). The hydrogen bonding in a β -pleated sheet takes place between different polypeptide chains, or different regions of a polypeptide chain that has turned back on itself. Each chain is fully extended, but because each has a zigzag structure, the resulting “sheet” has an overall pleated conformation (much like a sheet of paper that has been folded to make a fan). Although the pleated sheet is strong and flexible, it is not elastic. This is because the distance between the pleats is fixed, determined by the strong covalent bonds of the polypeptide backbones. Fibroin, the protein of silk, is characterized by a β -pleated sheet structure, as are the cores of many globular proteins.

It is not uncommon for a single polypeptide chain to include both α -helical regions and regions with β -pleated sheet

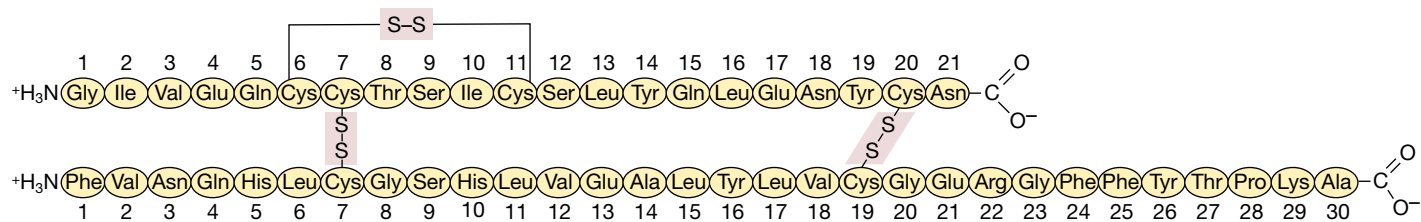
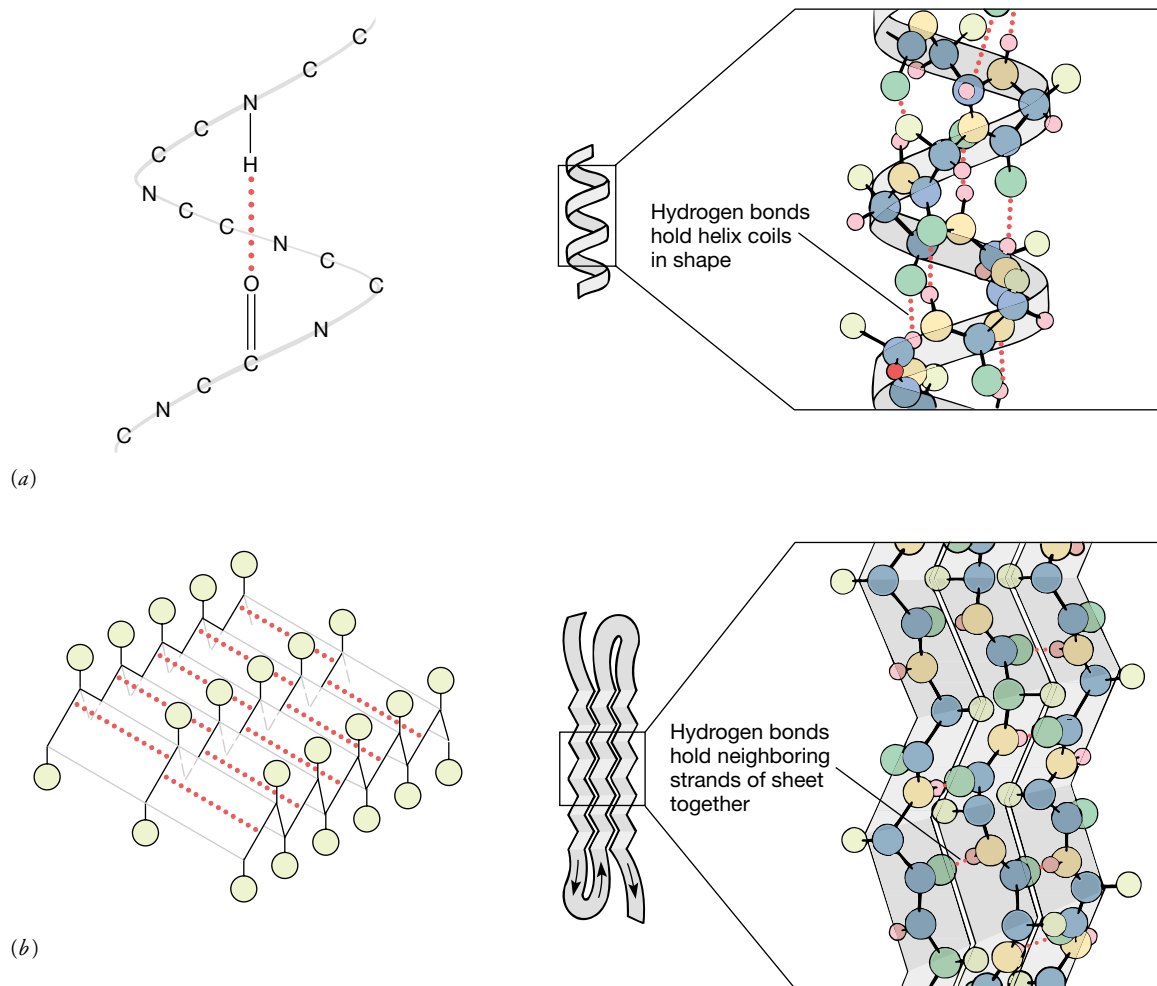


Figure 3-18 Primary structure of a polypeptide. Insulin is a very small protein made up of two polypeptides, each with its own primary structure. The linear sequence of amino acids is indicated by ovals containing their abbreviated names (see Fig. 3-14).

²Note that the designations α and β refer simply to the order in which these two types of secondary structures were discovered.



KEY: ● Carbon atom ● Oxygen atom ● Nitrogen atom ● Hydrogen atom ● R group



(c)

conformations. In addition, the properties of some complex biological materials result from such combinations. A spider's web is composed of a material that is extremely strong, flexible, and elastic. Once again we see function and structure

Figure 3–19 Secondary structure of a protein. (a) Note that the R groups project out from the sides of the α -helix. (The R groups have been omitted in the simplified diagram at left.) (b) In a β -pleated sheet, half the R groups project above the sheet and the other half project below it. (c) The strength and elasticity of a spider's web result from combining proteins with β -pleated sheet conformations and those with α -helical regions. (c, Skip Moody/Dembinsky Photo Associates)

working together, as these properties derive from the fact that it is a composite of proteins with α -helical conformations (providing elasticity) and others with β -pleated sheet conformations (providing strength) (Fig. 3–19c).

Tertiary structure depends on interactions among side chains

The **tertiary structure** of a protein molecule is the overall shape assumed by each individual polypeptide chain (Fig. 3–20). This three-dimensional structure is determined by four main factors that involve *interactions among R groups (side chains) belonging to the same polypeptide chain.*

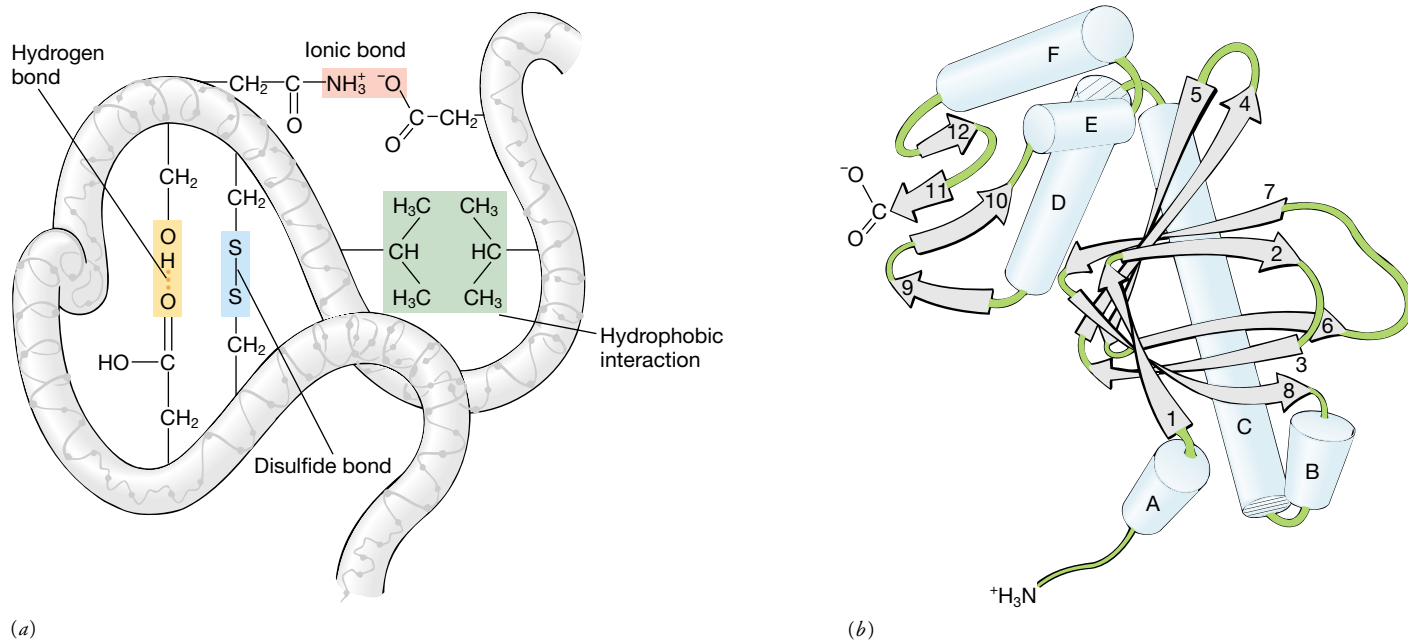


Figure 3–20 Tertiary structure of a protein. (a) Hydrogen bonds, hydrophobic interactions, and ionic attractions between R groups hold the parts of the molecule in the designated shape. Disulfide bonds are covalent bonds between the sulfur atoms of two cysteines. (b) Schematic drawing of the tertiary structure of a polypeptide. α -helical regions are represented as blue tubes lettered A through F; β -pleated sheets are the gray arrows numbered 1 through 12. Green lines represent connecting regions. Although the molecule seems very complicated, it is a single polypeptide chain, starting at the amino end (bottom left) and terminating at the carboxyl end. Most of the bends and foldbacks that give the molecule its overall conformation (tertiary structure) are stabilized by R-group interactions. This polypeptide is a subunit of a DNA-binding protein (CAP) from the bacterium *Escherichia coli*.

1. Hydrogen bonds form between R groups of certain amino acid subunits.
2. Ionic attraction can occur between an R group with a unit of positive charge and one with a unit of negative charge.
3. Hydrophobic interactions result from the tendency of non-polar R groups to be excluded by the surrounding water and therefore to associate in the interior of the globular structure.
4. Covalent bonds known as disulfide bonds or disulfide bridges (—S—S—) may link the sulfur atoms of two cysteine subunits belonging to the same chain. A disulfide bridge forms when the sulfhydryl groups of two cysteines react; the two hydrogens are removed and the two sulfur atoms that remain become covalently linked.

Quaternary structure results from interactions among polypeptides

Many functional proteins are composed of two or more polypeptide chains, interacting in specific ways to form the biologically active molecule. **Quaternary structure** is the resulting three-dimensional architecture of these polypeptide chains

(each with its own primary, secondary, and tertiary structure). The same types of interactions that produce secondary and tertiary structure can also contribute to quaternary structure; these include hydrogen bonding, ionic bonding, hydrophobic interactions, and disulfide bridges.

For example, Figure 3–18 illustrates not only primary structure, but also an aspect of quaternary structure; i.e., the two disulfide bridges joining the separate polypeptide chains that make up a functional insulin molecule. (Note that the disulfide bridges of insulin would contribute to tertiary structure if they were between cysteines belonging to the *same* polypeptide chain.) A functional antibody molecule consists of four polypeptide chains joined by disulfide bridges (Chapter 43). Disulfide bridges are a common feature of proteins, such as antibodies and insulin, that are secreted from cells; these strong bonds stabilize the molecules in the extracellular environment.

Hemoglobin, the protein in red blood cells that is responsible for oxygen transport, is an example of a globular protein with quaternary structure (Fig. 3–21). Hemoglobin consists of 574 amino acids arranged in four polypeptide chains: two identical chains called alpha chains and two identical chains called beta chains.

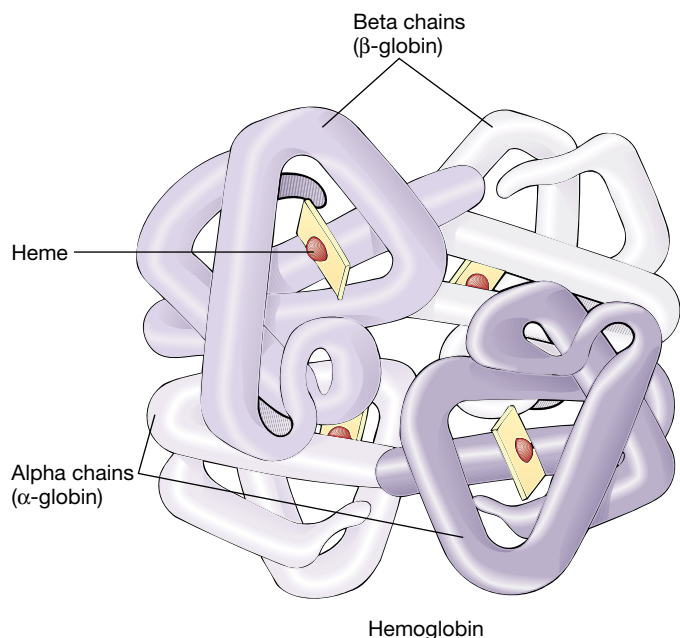


Figure 3–21 Quaternary structure of hemoglobin. Hemoglobin has quaternary structure because it consists of more than one polypeptide. Each polypeptide is joined to an iron-containing molecule, a heme.

The amino acid sequence of a protein determines its conformation

Under defined experimental conditions *in vitro* (that is, outside a living cell) a polypeptide can be demonstrated to spontaneously undergo folding processes that allow it to attain its normal, functional conformation. For example, researchers at the University of Illinois at Urbana-Champaign recently conducted an experiment in which they completely unfolded myoglobin, a polypeptide that stores oxygen in muscle cells, and then used sophisticated technology to track the refolding process. They found that within a few fractions of a microsecond the molecule had coiled up to form α -helices, and formation of the tertiary structure was completed within four microseconds.

This and other types of evidence support the widely held conclusion that amino acid sequence is the ultimate determinant of protein conformation. However, because conditions *in vivo* (in the cell) are quite different from defined laboratory conditions, proteins do not necessarily always fold spontaneously. On the contrary, in recent years it has been learned that proteins known as **molecular chaperones** mediate the folding of certain proteins. Chaperones are thought to make the folding process more orderly and efficient, and to prevent partially folded proteins from becoming inappropriately aggregated. However, there is no evidence that chaperones actually dictate the folding pattern. For this reason, the existence of chaperones is not an argument against the idea that amino acid sequence determines conformation.

Protein conformation is studied through a variety of methods

The conformation of a protein is ascertained directly through sophisticated types of analysis, such as the x-ray diffraction studies discussed in Chapter 11. Because these studies are tedious and costly, efforts have been made to develop alternative approaches. Today protein amino acid sequences can be determined rapidly through the application of genetic engineering techniques (Chapter 14). A variety of efforts are being made to effectively use these data to predict a protein's architecture. As we have seen, side chains can interact in relatively predictable ways, through ionic bonds, hydrogen bonds, etc. In addition, regions with certain types of side chains appear more likely to form α -helices or β -pleated sheets. Very complex computer programs are used to make such predictions, but these are imprecise because of the many possible combinations of folding patterns.

Computers are an essential part of yet another strategy. Once the amino acid sequence of a polypeptide has been determined, researchers use computers to search large databases, many accessible through the Internet, to find polypeptides with similar sequences. If the conformations of any of those polypeptides or portions of them are already known, this information can be extrapolated to make similar correlations between amino acid sequence and three-dimensional structure for the protein under investigation. These predictions are becoming increasingly reliable as more information is added to the databases on a daily basis.

Protein conformation determines function

The overall structure of a protein helps determine its biological activity. A single protein may have more than one distinct structural region, each with its own function. Many proteins are modular, consisting of two or more globular regions, called *domains*, connected by less compact regions of the polypeptide chain. Each domain may have a different function. For example, a protein might have one domain that attaches it to a membrane and another that allows it to act as an enzyme.

The biological activity of a protein can be disrupted by a change in amino acid sequence that results in a change in conformation. For example, the genetic disease known as *sickle cell anemia* (see Chapter 15) is due to a mutation that causes the substitution of the amino acid valine for glutamic acid at position 6 (the sixth amino acid from the amino end) in the beta chain of hemoglobin. The substitution of valine (which has a nonpolar side chain) for glutamic acid (which has a charged side chain) makes the hemoglobin less soluble and more likely to form crystal-like structures. This alteration of the hemoglobin affects the red blood cells, changing them to the crescent or sickle shapes that characterize this disease.

Changes in the three-dimensional structure of a protein also disrupt its biological activity. When a protein is heated, subjected to significant pH changes, or treated with any of a number of chemicals, its structure can become disordered and

the coiled peptide chains can unfold to give a more random conformation. This unfolding, which is mainly due to the disruption of hydrogen bonds and ionic bonds, is typically accompanied by a loss of normal function. Such changes in shape and the accompanying loss of biological activity are termed **denaturation** of the protein. For example, a denatured enzyme would lose its ability to catalyze a chemical reaction. An everyday example of denaturation occurs when we fry an egg. The consistency of the egg white protein, known as albumin, changes to a solid. Denaturation generally cannot be reversed (you can't "unfry" an egg). However, under certain conditions, some proteins have been denatured and have returned to their original shape and biological activity when normal environmental conditions have been restored.

DNA AND RNA ARE NUCLEIC ACIDS

Nucleic acids transmit hereditary information and determine what proteins a cell manufactures. There are two classes of nucleic acids found in cells: **ribonucleic acids (RNA)** and **deoxyribonucleic acids (DNA)**. DNA comprises the genes, the hereditary material of the cell, and contains instructions for making all the proteins, as well as all the RNA, needed by the organism. RNA is required as a direct participant in the complex process in which amino acids are linked to form polypeptides. Like proteins, nucleic acids are large, complex molecules. The name *nucleic acid* reflects the fact that they are acidic and were first identified, by Friederich Miescher in 1870, in the nuclei of pus cells.

Nucleic acids consist of nucleotide subunits

Nucleic acids are polymers of **nucleotides**, molecular units that consist of (1) a five-carbon sugar, either ribose or deoxyribose, (2) one or more phosphate groups, which make the molecule acidic, and (3) a nitrogenous base, a ring compound that contains nitrogen. The nitrogenous base may be either a double-ringed purine or a single-ringed pyrimidine (Fig. 3–22).

DNA commonly contains the purines adenine (A) and guanine (G), the pyrimidines cytosine (C) and thymine (T), the sugar deoxyribose, and phosphate. RNA contains the purines adenine and guanine, and the pyrimidines cytosine and uracil (U), together with the sugar ribose, and phosphate.

The molecules of nucleic acids are made of linear chains of nucleotides, which are joined by **phosphodiester linkages**, each consisting of a phosphate group and the covalent bonds that attach it to the sugars of adjacent nucleotides (Fig. 3–23). Note that each nucleotide is defined by its particular base and that nucleotides can be joined in any sequence. A nucleic acid molecule is uniquely defined by its specific sequence of nucleotides, which constitutes a kind of code (see Chapter 12). While RNA is usually composed of one nucleotide chain, DNA consists of two nucleotide chains held together by hydrogen bonds and entwined around each other in a double helix (see Fig. 1–7).

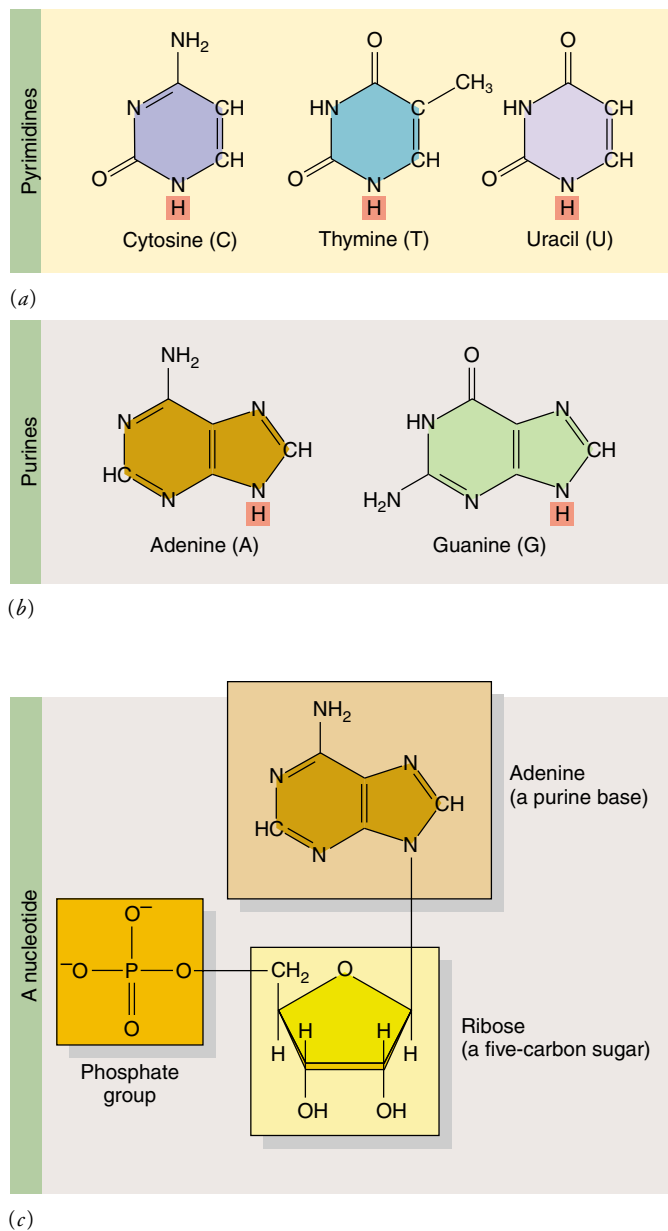


Figure 3–22 Nucleotides. (a) The three major pyrimidine bases found in nucleotides are cytosine, thymine (in DNA only), and uracil (in RNA only). (b) The two major purine bases found in nucleotides are adenine and guanine. The hydrogens indicated by the boxes are removed when the base is attached to a sugar. (c) A nucleotide, adenosine monophosphate (AMP).

Some nucleotides are important in energy transfers and other cellular functions

In addition to their importance as subunits of DNA and RNA, nucleotides serve other vital functions in living cells. **Adenosine triphosphate (ATP)**, composed of adenine, ribose, and three phosphates (Fig. 3–24), is of major importance as the primary energy currency of all cells (see Chapter 6). The two terminal phosphate groups are joined to the nucleotide by un-

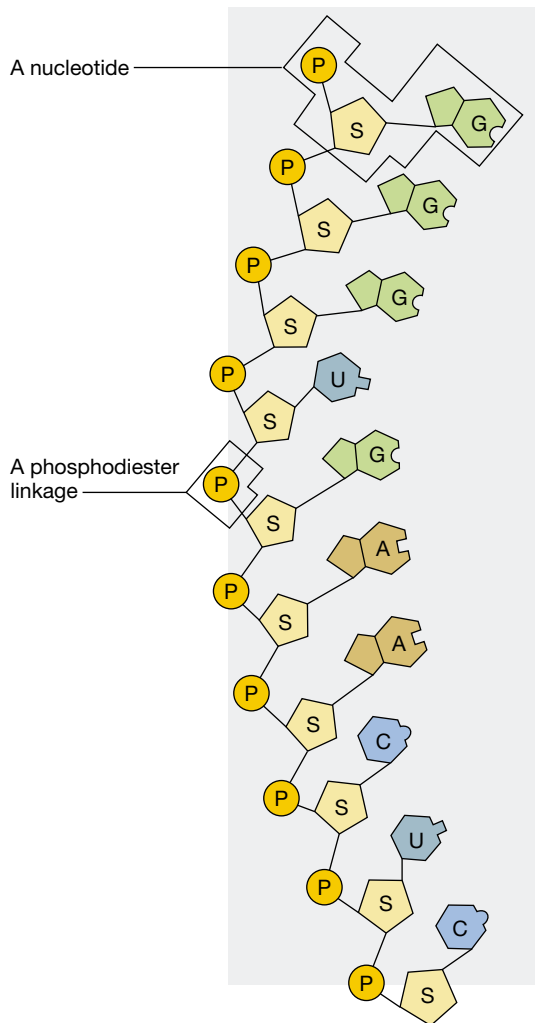


Figure 3–23 RNA. Nucleotides, each with a specific base, are joined by phosphodiester linkages. P, phosphate; S, the sugar ribose; G, guanine; C, cytosine; A, adenine; U, uracil.

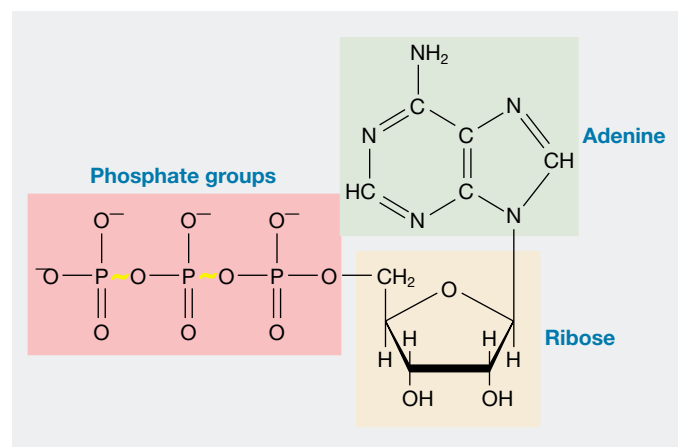


Figure 3–24 ATP, a nucleotide. The two terminal phosphate groups are joined by unstable bonds (indicated by wavy lines). These bonds permit the phosphates to be transferred to other molecules, making them more reactive.

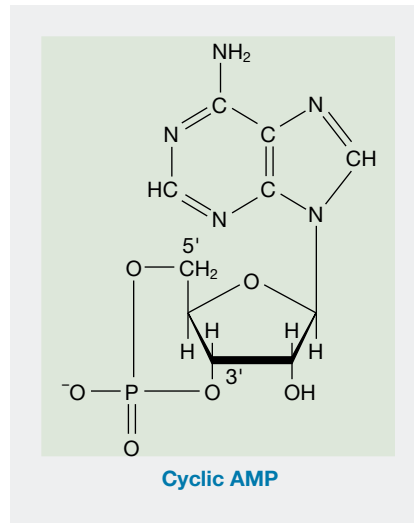


Figure 3–25 Cyclic AMP. The single phosphate is part of a ring connecting two regions of the ribose.

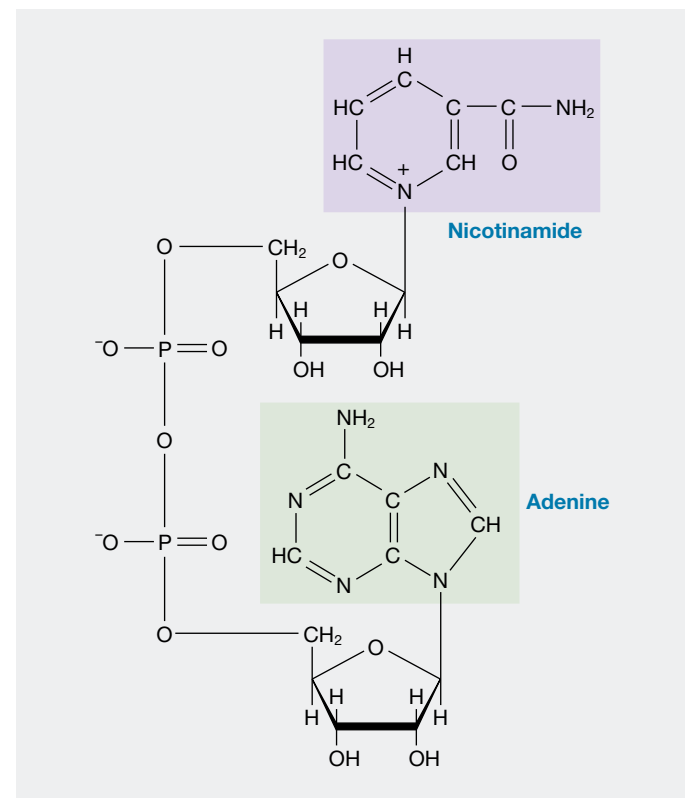


Figure 3–26 NAD⁺, an important hydrogen (electron) acceptor. The nicotinamide portion of the molecule accepts hydrogen and becomes reduced in the process. The resulting reduced molecule, known as NADH, is an electron donor.

stable covalent bonds, traditionally indicated by wavy lines. ATP can transfer a phosphate group to another molecule, making that molecule more reactive. In this way ATP is able to donate some of its chemical energy. Most of the readily available chemical energy of the cell is associated with the phosphate groups of ATP.

A nucleotide may be converted to an alternative form with specific cellular functions. ATP, for example, is converted to cyclic AMP (c-AMP or cyclic adenosine monophosphate) by the enzyme adenyl cyclase (Fig. 3–25). Cyclic AMP regulates certain cellular functions and is important in the mechanism by which some hormones act (see Chapters 13, 39, and 47).

Cells contain several dinucleotides, which are of great importance in metabolic processes. For example, as discussed in Chapter 6, **nicotinamide adenine dinucleotide** (Fig. 3–26)

has a primary role in biological oxidations and reductions within cells. It can exist in an oxidized form (**NAD⁺**) that is converted to a reduced form (**NADH**) when it accepts electrons (in association with hydrogen). These electrons, along with their energy, can be transferred to other molecules.

BIOLOGICAL MOLECULES CAN BE RECOGNIZED BY THEIR KEY FEATURES

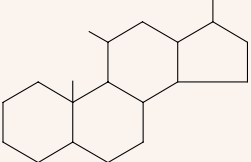
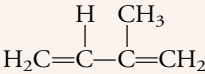
Although the fundamental classes of biological molecules may seem to form a bewildering array, one can learn to distinguish them readily by understanding their chief attributes. These are summarized in Table 3–2.

TABLE 3–2 Some of the Groups of Biologically Important Organic Compounds

Class of compounds	Component elements	Description	How to recognize	Principal function in living systems
Carbohydrates	C, H, O	Contain approximately 1 C:2 H:1 O (but make allowance for loss of oxygen and hydrogen when sugar units are linked)	Count the carbons, hydrogens, and oxygens.	Cellular fuel; energy storage; structural component of plant cell walls; component of other compounds such as nucleic acids and glycoproteins
		1. Monosaccharides (simple sugars) Mainly five-carbon (pentose) molecules such as ribose or six-carbon (hexose) molecules such as glucose and fructose	Look for the ring shapes:	Cellular fuel; components of other compounds
		2. Disaccharides. Two sugar units linked by a glycosidic bond, e.g., maltose, sucrose	Count sugar units	Components of other compounds; form of sugar transported in plants
		3. Polysaccharides. Many sugar units linked by glycosidic bonds, e.g., glycogen, cellulose	Count sugar units	Energy storage; structural components of plant cell walls
Lipids	C, H, O	Contain much less oxygen relative to carbon and hydrogen than do carbohydrates		Energy storage; cellular fuel, structural components of cells; thermal insulation
		1. Neutral fats. Combination of glycerol with one to three fatty acids. Monoacylglycerol contains one fatty acid; diacylglycerol contains two fatty acids; triacylglycerol contains three fatty acids. If fatty acids contain double carbon-to-carbon linkages (C=C), they are unsaturated; otherwise they are saturated	Look for glycerol at one end of molecule:	Cellular fuel; energy storage

continued

TABLE 3-2 continued

Class of compounds	Component elements	Description	How to recognize	Principal function in living systems
		2. Phospholipids. Composed of glycerol attached to one or two fatty acids and to an organic base containing phosphorus	Look for glycerol and side chain containing phosphorus and nitrogen.	Components of cell membranes
		3. Steroids. Complex molecules containing carbon atoms arranged in four attached rings. (Three rings contain six carbon atoms each, and the fourth ring contains five.)	Look for four attached rings: 	Some are hormones, others include cholesterol, bile salts, vitamin D, components of cell membranes
		4. Carotenoids. Orange and yellow pigments; consist of isoprene units	Look for isoprene units. 	Retinal (important in photoreception) and vitamin A are formed from carotenoids.
Proteins	C, H, O, N, usually S	One or more polypeptides (chains of amino acids) coiled or folded in characteristic shapes	Look for amino acid units joined by C—N bonds.	Serve as enzymes; structural components; muscle proteins; hemoglobin
Nucleic acids	C, H, O, N, P	Backbone composed of alternating pentose and phosphate groups, from which nitrogenous bases project. DNA contains the sugar deoxyribose and the bases guanine, cytosine, adenine, and thymine. RNA contains the sugar ribose and the bases guanine, cytosine, adenine, and uracil. Each molecular subunit, called a <i>nucleotide</i> , consists of a pentose, a phosphate, and a nitrogenous base.	Look for a pentose-phosphate backbone. DNA forms a double helix.	Storage, transmission, and expression of genetic information

SUMMARY WITH KEY TERMS

- I. The major groups of biologically important organic compounds are carbohydrates, lipids, proteins, and nucleic acids.
- II. The properties of carbon atoms make them extraordinarily versatile, able to form the backbones of the large variety of organic compounds essential to life.
 - A. Each carbon atom can form four covalent bonds with four other atoms; these can be single, double, or triple bonds.
 - B. Carbon forms covalent bonds with a greater number of different elements than does any other type of atom. Carbon atoms can form straight or branched chains or can join into rings.
- III. **Isomers** are compounds with the same molecular formula but different structures.
 - A. **Structural isomers** differ in the covalent arrangements of their atoms.
 - B. **Geometric isomers**, or *cis-trans* isomers, differ in the spatial arrangements of their atoms.
 - C. **Enantiomers** are isomers that are mirror images of each other. Cells can distinguish between these configurations.
- IV. Organic compounds are made up of specific functional groups with characteristic properties.
 - A. **Hydrocarbons** are nonpolar and hydrophobic.
 - B. Polar and ionic functional groups interact with each other and dissolve in water.
 - C. Partial charges on atoms at opposite ends of a bond are responsible for the polar property of a functional group. **Hydroxyl** and **carbonyl** groups are polar.
 - D. **Carboxyl** and **phosphate** groups are acidic, becoming negatively charged when they release hydrogen ions. The **amino** group is basic, becoming positively charged when it accepts a hydrogen ion.
- V. Long chains of similar organic compounds linked together are called **polymers**. Large polymers such as polysaccharides, proteins, and DNA are referred to as **macromolecules**.
- VI. **Carbohydrates** contain carbon, hydrogen, and oxygen in a ratio of approximately one carbon to two hydrogens to one oxygen.
 - A. **Monosaccharides** are simple sugars such as glucose, fructose, and ribose.

- B. Two monosaccharides can be joined by a glycosidic linkage, forming a **disaccharide** such as maltose or sucrose.
- C. Most carbohydrates are **polysaccharides**, long chains of repeating units of a simple sugar.
- Carbohydrates are typically stored in plants as **starch** and in animals as **glycogen**.
 - The cell walls of plants are composed mainly of the polysaccharide **cellulose**.
- VII. **Lipid** molecules are composed mainly of hydrocarbon-containing regions, with few oxygen-containing (polar or ionic) functional groups. Lipids have a greasy or oily consistency and are relatively insoluble in water.
- A. **Neutral fats** are used for fuel storage. A fat consists of a molecule of **glycerol** combined with one to three **fatty acids**.
- Monoacylglycerols**, **diacylglycerols**, and **triacylglycerols** are neutral fats containing one, two, and three fatty acids, respectively.
 - A fatty acid can be either **saturated** with hydrogen, or **unsaturated**.
- B. **Phospholipids** are structural components of cellular membranes.
- C. **Steroid** molecules contain carbon atoms arranged in four attached rings. Cholesterol, bile salts, and certain hormones are important steroids.
- VIII. **Proteins** are large, complex molecules made of simpler subunits, called **amino acids**, joined by **peptide bonds**.
- A. Proteins are the most versatile class of biological molecules, serving as **enzymes**, structural components, cellular regulators, etc.
- B. Proteins are composed of various linear sequences of 20 different amino acids. Two amino acids combine to form a **dipeptide**. A longer chain of amino acids is a **polypeptide**.
- All amino acids contain an amino group and a carboxyl group, but vary in their side chains. The side chains of amino acids dictate their chemical properties.
 - Amino acids generally exist as dipolar ions at cellular pH and serve as important biological buffers.
- C. Four levels of organization can be distinguished in protein molecules.
- Primary structure** is the linear sequence of amino acids in the polypeptide chain.
 - Secondary structure** is a regular conformation, such as an **α -helix** or a **β -pleated sheet**, due to hydrogen bonding between elements of the uniform backbone of the polypeptide.
 - Tertiary structure** is the overall shape of the polypeptide chains, as dictated by chemical properties and interactions of the side chains of specific amino acids. Hydrogen bonding, ionic bonds, hydrophobic interactions, and disulfide bridges contribute to tertiary structure.
 - Quaternary structure** is determined by the association of two or more polypeptide chains.
- IX. The **nucleic acids** DNA and RNA store and transfer information that governs the sequence of amino acids in proteins and ultimately the structure and function of the organism.
- A. Nucleic acids are composed of long chains of **nucleotide** subunits, each composed of a purine or pyrimidine nitrogenous base, a five-carbon sugar (ribose or deoxyribose), and a phosphate group.
- B. **ATP** (adenosine triphosphate) is a nucleotide of special significance in energy metabolism. **NAD⁺** is also involved in energy metabolism through its role as an electron (hydrogen) acceptor in biological oxidations.

POST - TEST

- Which of the following would be considered to be an inorganic form of carbon? (a) H_2CO_3 (b) C_2H_4 (c) CH_3COOH (d) b and c (e) all of the above are inorganic
- Carbon is particularly well suited to be the backbone of organic molecules because (a) it can form both covalent bonds and ionic bonds (b) its covalent bonds are very irregularly arranged in three-dimensional space (c) its covalent bonds are the strongest chemical bonds known (d) it can bond to a large number of other types of atoms (e) all of the bonds it forms are polar
- The structures depicted below are

$$\begin{array}{c} \text{CH}_3 \text{CH}_3 \\ | \quad | \\ \text{H}-\text{C}-\text{C}-\text{H} \\ | \quad | \\ \text{H} \quad \text{H} \end{array}$$

$$\begin{array}{c} \text{H} \quad \text{CH}_3 \\ | \quad | \\ \text{H}-\text{C}-\text{C}-\text{H} \\ | \quad | \\ \text{CH}_3 \text{H} \end{array}$$

(a) enantiomers (b) different views of the same molecule (c) geometric (*cis-trans*) isomers (d) both geometric isomers and enantiomers (e) structural isomers
- Which of the following are generally hydrophobic? (a) polar molecules and hydrocarbons (b) ions and hydrocarbons (c) nonpolar molecules and ions (d) polar molecules and ions (e) none of the above
- Which of the following is a nonpolar molecule? (a) water (H_2O) (b) ammonia (NH_3) (c) methane (CH_4) (d) ethane (C_2H_6) (e) more than one of the above
- Which of the following functional groups normally acts as an acid? (a) hydroxyl (b) carbonyl (c) sulfhydryl (d) phosphate (e) amino
- A monosaccharide designated as an aldehyde sugar contains (a) a terminal carboxyl group (b) an internal carboxyl group (c) a terminal carbonyl group (d) an internal carbonyl group (e) a terminal carboxyl group and an internal carbonyl group
- Structural polysaccharides typically (a) have extensive hydrogen bonding between adjacent molecules (b) are much more hydrophilic than storage polysaccharides (c) have much stronger covalent bonds than do storage polysaccharides (d) consist of alternating α -glucose and β -glucose subunits (e) form helical structures in the cell
- Fatty acids are components of (a) phospholipids and carotenoids (b) carotenoids and triacylglycerol (c) steroids and triacylglycerol (d) phospholipids and triacylglycerol (e) carotenoids and steroids
- Saturated fatty acids are so named because they are saturated with (a) hydrogen (b) water (c) hydroxyl groups (d) glycerol (e) double bonds
- Which pair of amino acid side groups would be most likely to associate with each other by an ionic bond?
 - $-\text{CH}_3$
 - $-\text{CH}_2-\text{COO}^-$
 - $-\text{CH}_2-\text{CH}_2-\text{NH}_3^+$
 - $-\text{CH}_2-\text{CH}_2-\text{COO}^-$
 - $-\text{CH}_2-\text{OH}$

(a) 1 and 2 (b) 2 and 4 (c) 1 and 5 (d) 2 and 5 (e) 3 and 4
- Which of the following levels of protein structure may be affected by hydrogen bonding? (a) primary and secondary (b) primary and tertiary (c) secondary, tertiary, and quaternary (d) primary, secondary, and tertiary (e) primary, secondary, tertiary, and quaternary
- Each phosphodiester linkage in DNA or RNA includes a phosphate joined by covalent bonds to (a) two bases (b) two sugars (c) two additional phosphates (d) a sugar, a base, and a phosphate (e) a sugar and a base

REVIEW QUESTIONS

1. Marble is composed of the carbon-containing compound calcium carbonate (CaCO_3). Should calcium carbonate be considered an organic molecule? Why or why not?
2. What are some of the ways that the features of carbon-to-carbon bonds influence the stability and three-dimensional structure of organic molecules?
3. Draw pairs of simple sketches comparing two (a) structural isomers, (b) geometric isomers, and (c) enantiomers.
4. Sketch the following functional groups: methyl, amino, carbonyl, hydroxyl, carboxyl, and phosphate. Classify each as either nonpolar, polar, acidic, or basic.
5. What features related to hydrogen bonding give storage polysaccharides, such as starch and glycogen, very different properties from structural polysaccharides, such as cellulose and chitin?
6. Draw a structural formula of a simple amino acid and identify the carboxyl group, amino group, and R group.
7. How does the primary structure of a polypeptide influence its secondary and tertiary structures? How can the conformation of a protein be disrupted?
8. Compare the functions of proteins and nucleic acids.

YOU MAKE THE CONNECTION

1. Like oxygen, sulfur forms two covalent bonds. However, it is far less electronegative. In fact, it is approximately as electronegative as carbon. How would the properties of the various classes of biological molecules be altered if you were to replace all the oxygen atoms with sulfur atoms?
2. How would you reconcile the following statements? (1) All organisms share a common ancestry. (2) Individual organisms are generally biochemically unique.

RECOMMENDED READINGS

Atkins, P.W. *Molecules*. Scientific American Library, W.H. Freeman and Co., New York, 1987. A fascinating account of the relationship between the properties and the molecular structures of various substances.

Bettelheim, F.A. and J. March. *Introduction to General, Organic and Biochemistry*, 4th ed. Saunders College Publishing, Philadelphia, 1995. A very readable reference text for those who would like to know more about the chemistry basic to life.

Garrett, R.H. and C.M. Grisham. *Biochemistry*. Saunders College Publishing, Philadelphia, 1995. A comprehensive, advanced biochemistry text.

Richards, F.M. "The Protein Folding Problem." *Scientific American*, Vol. 264, No. 1, Jan. 1991. A discussion of the mechanisms involved when a protein folds into its biologically active shape.

● Visit our website at <http://www.saunderscollege.com/lifesci/titles.html> and click on Solomon/Berg/Martin Biology for links to chapter-related resources on the World Wide Web.

