7

Carbohydrates and Glycobiology

Carbohydrates are the most abundant biomolecules on Earth. Each year, photosynthesis converts more than 100 billion metric tons of CO_2 and H_2O into cellulose and other plant products. Certain carbohydrates (sugar and starch) are a dietary staple in most parts of the world, and the oxidation of carbohydrates is the central energy-yielding pathway in most nonphotosynthetic cells. Carbohydrate polymers (also called glycans) serve as structural and protective elements in the cell walls of bacteria and plants and in the connective tissues of animals. Other carbohydrate polymers lubricate skeletal joints and participate in recognition and adhesion between cells. Complex carbohydrate polymers covalently attached to proteins or lipids act as signals that determine the intracellular destination or metabolic fate of these hybrid molecules, called **glycoconjugates**. This chapter introduces the major classes of carbohydrates and glycoconjugates and provides a few examples of their many structural and functional roles.

Carbohydrates are polyhydroxy aldehydes or ketones, or substances that yield such compounds on hydrolysis. Many, but not all, carbohydrates have the empirical formula $(CH_2O)_n$; some also contain nitrogen, phosphorus, or sulfur. There are three major size classes of carbohydrates: monosaccharides, oligosaccharides, and polysaccharides (the word "saccharide" is derived from the Greek *sakcharon*, meaning "sugar"). **Monosaccharides**, or simple sugars, consist of a single polyhydroxy aldehyde or ketone unit. The most abundant monosaccharide in nature is the six-carbon sugar D-glucose, sometimes referred to as dextrose. Monosaccharides of four or more carbons tend to have cyclic structures.

Oligosaccharides consist of short chains of monosaccharide units, or residues, joined by characteristic linkages called glycosidic bonds. The most abundant are the **disaccharides**, with two monosaccharide units. Typical is sucrose (cane sugar), which consists of the six-carbon sugars D-glucose and D-fructose. All common monosaccharides and disaccharides have names ending with the suffix "-ose." In cells, most oligosaccharides consisting of three or more units do not occur as free entities but are joined to nonsugar molecules (lipids or proteins) in glycoconjugates.

The **polysaccharides** are sugar polymers containing more than 20 or so monosaccharide units; some have hundreds or thousands of units. Some polysaccharides, such as cellulose, are linear chains; others, such as glycogen, are branched. Both glycogen and cellulose consist of recurring units of D-glucose, but they differ in the type of glycosidic linkage and consequently have strikingly different properties and biological roles.

7.1 Monosaccharides and Disaccharides

The simplest of the carbohydrates, the monosaccharides, are either aldehydes or ketones with two or more hydroxyl groups; the six-carbon monosaccharides glucose and fructose have five hydroxyl groups. Many of the carbon atoms to which hydroxyl groups are attached are chiral centers, which give rise to the many sugar stereoisomers found in nature. Stereoisomerism in sugars is biologically significant because the enzymes that act on sugars are strictly stereospecific, typically preferring one stereoisomer to another by three or more orders of magnitude, as reflected in K_m values or binding constants. It is as difficult to fit the wrong sugar stereoisomer into an enzyme's binding site as it is to put your left glove on your right hand.

We begin by describing the families of monosaccharides with backbones of three to seven carbons—their structure and stereoisomeric forms, and the means of representing their three-dimensional structures on paper. We then discuss several chemical reactions of the carbonyl groups of monosaccharides. One such reaction, the addition of a hydroxyl group from within the same molecule, generates cyclic forms having four or more backbone carbons (the forms that predominate in aqueous solution). This ring closure creates a new chiral center, adding further stereochemical complexity to this class of compounds. The nomenclature for unambiguously specifying the configuration about each carbon atom in a cyclic form and the means of representing these structures on paper are therefore described in some detail; this information will be useful as we discuss the metabolism of monosaccharides in Part II. We also introduce here some important monosaccharide derivatives encountered in later chapters.

The Two Families of Monosaccharides Are Aldoses and Ketoses

Monosaccharides are colorless, crystalline solids that are freely soluble in water but insoluble in nonpolar solvents. Most have a sweet taste (see Box 7–2, p. 254). The backbones of common monosaccharides are unbranched carbon chains in which all the carbon atoms are linked by single bonds. In this open-chain form, one of the carbon atoms is double-bonded to an oxygen atom to form a carbonyl group; each of the other carbon atoms has a hydroxyl group. If the carbonyl group is at an end of the carbon chain (that is, in an aldehyde group) the monosaccharide is an **aldose**; if the carbonyl group is at any other position (in a ketone group) the monosaccharide is a **ketose**. The simplest monosaccharides are the two three-carbon trioses: glyceraldehyde, an aldotriose, and dihydroxyacetone, a ketotriose **(Fig. 7–1a)**.

Monosaccharides with four, five, six, and seven carbon atoms in their backbones are called, respectively, tetroses, pentoses, hexoses, and heptoses. There are aldoses and ketoses of each of these chain lengths: aldotetroses and ketotetroses, aldopentoses and ketopentoses, and so on. The hexoses, which include the aldohexose D-glucose and the ketohexose D-fructose (Fig. 7–1b), are the most common monosaccharides in nature—the products of photosynthesis, and key intermediates in the central energy-yielding reaction sequence in most organisms. The aldopentoses D-ribose and 2-deoxy-D-ribose (Fig. 7–1c) are components of nucleotides and nucleic acids (Chapter 8).

Monosaccharides Have Asymmetric Centers

All the monosaccharides except dihydroxyacetone contain one or more asymmetric (chiral) carbon atoms and thus occur in optically active isomeric forms (pp. 17–18). The simplest aldose, glyceraldehyde, contains one chiral center (the middle carbon atom) and therefore has two different optical isomers, or **enantiomers (Fig. 7–2)**.

KEY CONVENTION: One of the two enantiomers of glyceraldehyde is, by convention, designated the D isomer, the other the L isomer. As for other biomolecules with chiral centers, the absolute configurations of sugars are known from x-ray crystallography. To represent three-dimensional sugar structures on paper, we often use **Fischer projection formulas** (Fig. 7–2). In Fischer projection formulas, horizontal bonds project out of the plane of the paper, toward the reader; vertical bonds project behind the plane of the paper, away from the reader. ■

In general, a molecule with n chiral centers can have 2^n stereoisomers. Glyceraldehyde has $2^1 = 2$; the aldohexoses, with four chiral centers, have $2^4 = 16$. The stereoisomers of monosaccharides of each carbon-chain

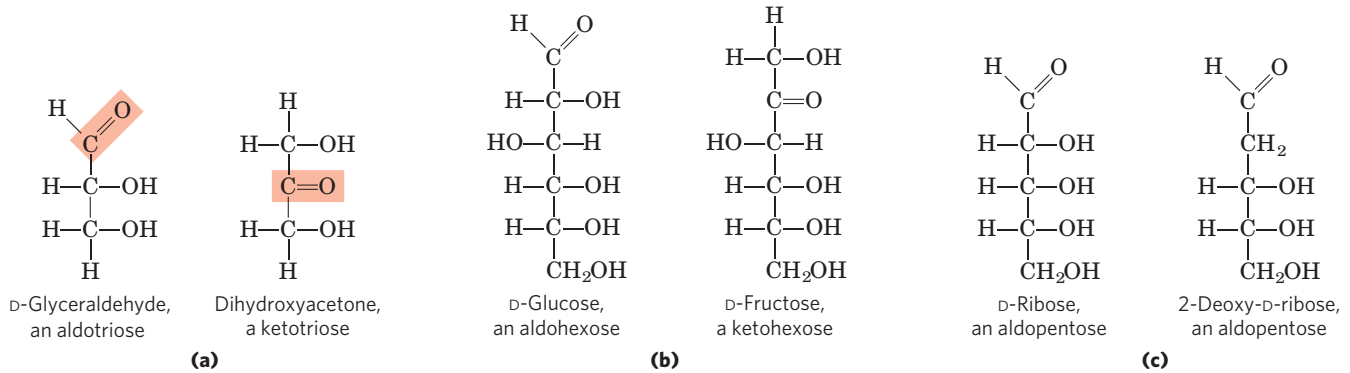


FIGURE 7–1 Representative monosaccharides. **(a)** Two trioses, an aldose and a ketose. The carbonyl group in each is shaded. **(b)** Two common hexoses. **(c)** The pentose components of nucleic acids. D-Ribose is a component of ribonucleic acid (RNA), and 2-deoxy-D-ribose is a component of deoxyribonucleic acid (DNA).

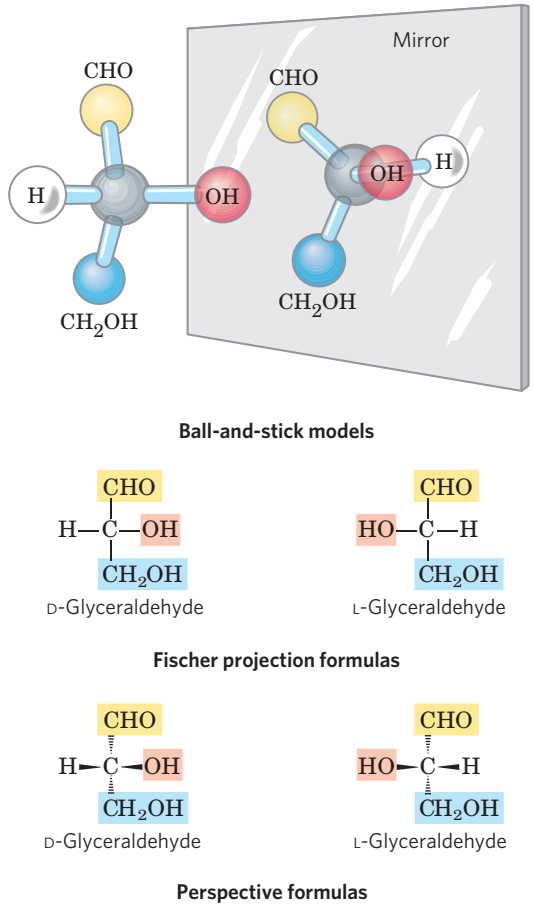
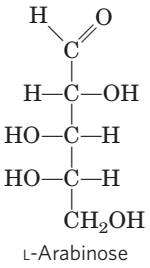


FIGURE 7–2 Three ways to represent the two enantiomers of glyceraldehyde. The enantiomers are mirror images of each other. Ball-and-stick models show the actual configuration of molecules. Recall (see Fig. 1–18) that in perspective formulas, the wide end of a solid wedge projects out of the plane of the paper, toward the reader; a dashed wedge extends behind.

length can be divided into two groups that differ in the configuration about the chiral center *most distant* from the carbonyl carbon. Those in which the configuration at this reference carbon is the same as that of D-glyceraldehyde are designated D isomers, and those with the same configuration as L-glyceraldehyde are L isomers. In other words, when the hydroxyl group on the reference carbon is on the right (*dextro*) in a projection formula that has the carbonyl carbon at the top, the sugar is the D isomer; when on the left (*levo*), it is the L isomer. Of the 16 possible aldohexoses, eight are D forms and eight are L. Most of the hexoses of living organisms are D isomers. Why D isomers? An interesting and unanswered question. Recall that all of the amino acids found in protein are exclusively one of two possible stereoisomers, L. The basis for this initial preference for one isomer during evolution is also unknown; however, once one isomer had been selected, it was likely that evolving enzymes would retain their preference for that stereoisomer (p. 78).

Figure 7–3 shows the structures of the D stereoisomers of all the aldoses and ketoses having three to six carbon atoms. The carbons of a sugar are numbered beginning at the end of the chain nearest the carbonyl group. Each of the eight D-aldohexoses, which differ in the stereochemistry at C-2, C-3, or C-4, has its own name: D-glucose, D-galactose, D-mannose, and so forth (Fig. 7–3a). The four- and five-carbon ketoses are designated by inserting "ul" into the name of a corresponding aldose; for example, D-ribulose is the ketopentose corresponding to the aldopentose D-ribose. (We will see the importance of ribulose when we discuss the fixation of atmospheric CO_2 by green plants, in Chapter 20.) The ketohexoses are named otherwise: for example, fructose (from the Latin *fructus*, "fruit"; fruits are one source of this sugar) and sorbose (from *Sorbus*, the genus of mountain ash, which has berries rich in the related sugar alcohol sorbitol). Two sugars that differ only in the configuration around one carbon atom are called **epimers**; D-glucose and D-mannose, which differ only in the stereochemistry at C-2, are epimers, as are D-glucose and D-galactose (which differ at C-4) **(Fig. 7–4)**.

Some sugars occur naturally in their L form; examples are L-arabinose and the L isomers of some sugar derivatives that are common components of glycoconjugates (Section 7.3).



L-Arabinose

The Common Monosaccharides Have Cyclic Structures

For simplicity, we have thus far represented the structures of aldoses and ketoses as straight-chain molecules (Figs 7–3, 7–4). In fact, in aqueous solution, aldotetroses and all monosaccharides with five or more carbon atoms in the backbone occur predominantly as cyclic (ring) structures in which the carbonyl group has formed a covalent bond with the oxygen of a hydroxyl group along the chain. The formation of these ring structures is the result of a general reaction between alcohols and aldehydes or ketones to form derivatives called **hemiacetals** or **hemiketals**. Two molecules of an alcohol can add to a carbonyl carbon; the product of the first addition is a hemiacetal (for addition to an aldose) or a hemiketal (for addition to a ketose). If the —OH and carbonyl groups are from the same molecule,

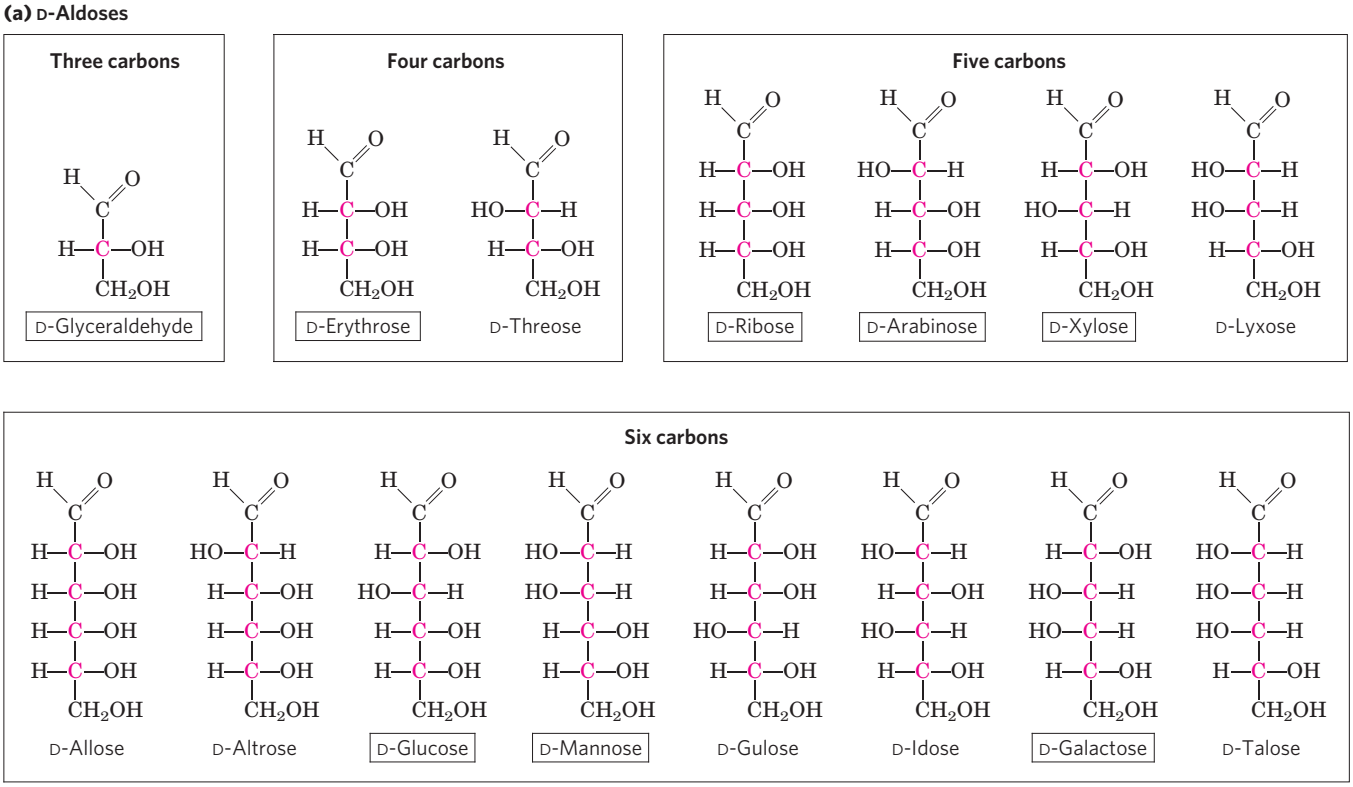


FIGURE 7–3 Aldoses and ketoses. The series of **(a)** D-aldoses and **(b)** D-ketoses having from three to six carbon atoms, shown as projection formulas. The carbon atoms in red are chiral centers. In all these D isomers, the chiral carbon *most distant from the carbonyl carbon* has the same configuration as the chiral carbon in D-glyceraldehyde. The sugars named in boxes are the most common in nature; you will encounter these again in this and later chapters.

a five- or six-membered ring results. Addition of the second molecule of alcohol produces the full acetal or ketal **(Fig. 7–5)**, and the bond formed is a glycosidic linkage. When the two molecules that react are both monosaccharides, the acetal or ketal formed is a disaccharide.

The reaction with the first molecule of alcohol creates an additional chiral center (the carbonyl carbon). Because the alcohol can add in either of two ways, attacking either the "front" or the "back" of the carbonyl carbon, the reaction can produce either of two stereoisomeric configurations, denoted α and β. For example, D-glucose exists in solution as an intramolecular hemiacetal in which the free hydroxyl group at C-5 has reacted with the aldehydic C-1, rendering the latter carbon asymmetric and producing two possible stereoisomers, designated α and β **(Fig. 7–6)**. Isomeric forms of monosaccharides that differ only in their configuration about the hemiacetal or hemiketal carbon atom are called **anomers**, and the carbonyl carbon atom is called the **anomeric carbon**.

Six-membered ring compounds are called **pyranoses** because they resemble the six-membered ring compound

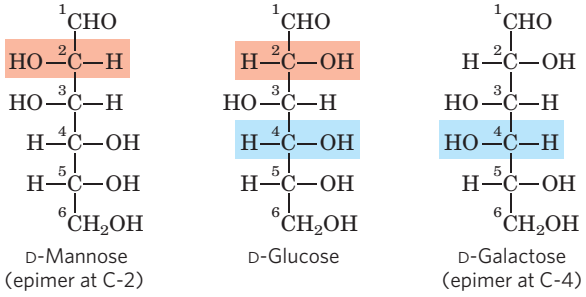


FIGURE 7–4 Epimers. D-Glucose and two of its epimers are shown as projection formulas. Each epimer differs from D-glucose in the configuration at one chiral center (shaded light red or blue).

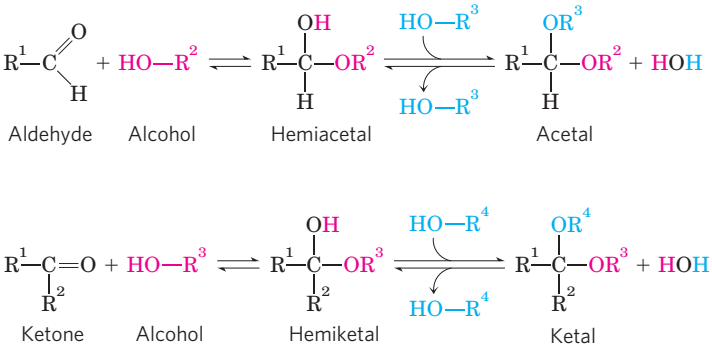


FIGURE 7–5 Formation of hemiacetals and hemiketals. An aldehyde or ketone can react with an alcohol in a 1:1 ratio to yield a hemiacetal or hemiketal, respectively, creating a new chiral center at the carbonyl carbon. Substitution of a second alcohol molecule produces an acetal or ketal. When the second alcohol is part of another sugar molecule, the bond produced is a glycosidic bond (p. 252).

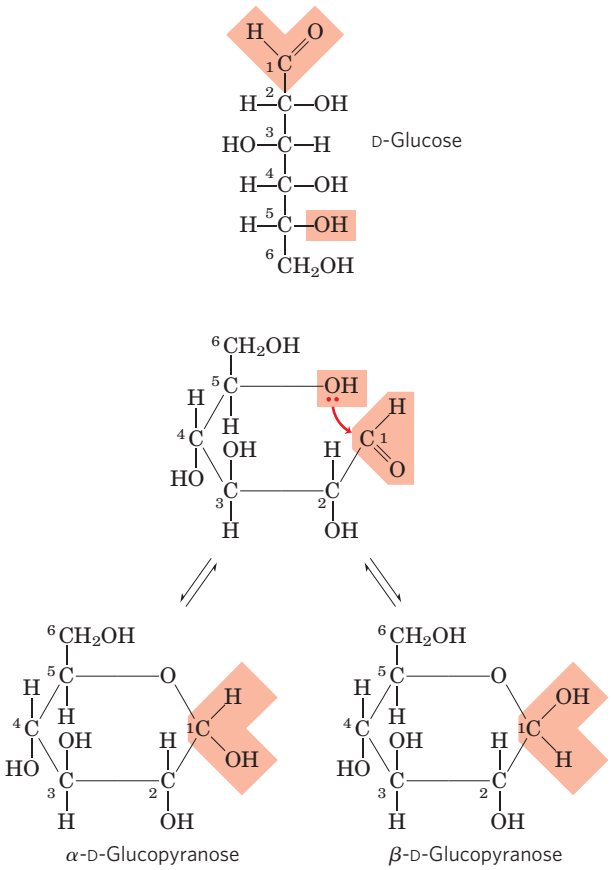
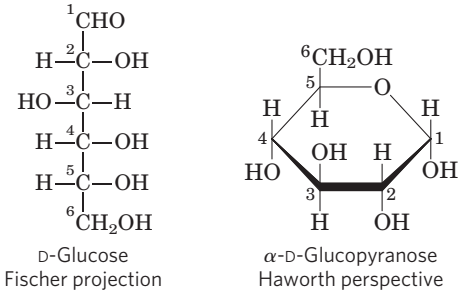


FIGURE 7–6 Formation of the two cyclic forms of D-glucose. Reaction between the aldehyde group at C-1 and the hydroxyl group at C-5 forms a hemiacetal linkage, producing either of two stereoisomers, the α and β anomers, which differ only in the stereochemistry around the hemiacetal carbon. This reaction is reversible. The interconversion of α and β anomers is called mutarotation.

pyran **(Fig. 7–7)**. The systematic names for the two ring forms of D-glucose are therefore α-D-glucopyranose and β-D-glucopyranose. Ketohexoses (such as fructose) also occur as cyclic compounds with α and β anomeric forms. In these compounds the hydroxyl group at C-5 (or C-6) reacts with the keto group at C-2, forming a **furanose** (or pyranose) ring containing a hemiketal linkage (Fig. 7–5). D-Fructose readily forms the furanose ring (Fig. 7–7); the more common anomer of this sugar in combined forms or in derivatives is β-D-fructofuranose.

Cyclic sugar structures are more accurately represented in **Haworth perspective formulas** than in the Fischer projections commonly used for linear sugar structures. In Haworth projections the six-membered ring is tilted to make its plane almost perpendicular to that of the paper, with the bonds closest to the reader drawn thicker than those farther away, as in Figure 7–7.

KEY CONVENTION: To convert the Fischer projection formula of any linear D-hexose to a Haworth perspective formula showing the molecule's cyclic structure, draw the six-membered ring (five carbons and one oxygen, at the upper right), number the carbons in a clockwise direction beginning with the anomeric carbon, then place the hydroxyl groups. If a hydroxyl group is to the right in the Fischer projection, it is placed pointing down (i.e., below the plane of the ring) in the Haworth perspective; if it is to the left in the Fischer projection, it is placed pointing up (i.e., above the plane) in the Haworth perspective. The terminal —CH_2OH group projects upward for the D-enantiomer, downward for the L-enantiomer. The hydroxyl on the anomeric carbon can point up or down. When the anomeric hydroxyl of a D-hexose is on the same side of the ring as C-6, the structure is by definition β; when it is on the opposite side from C-6, the structure is α. ■



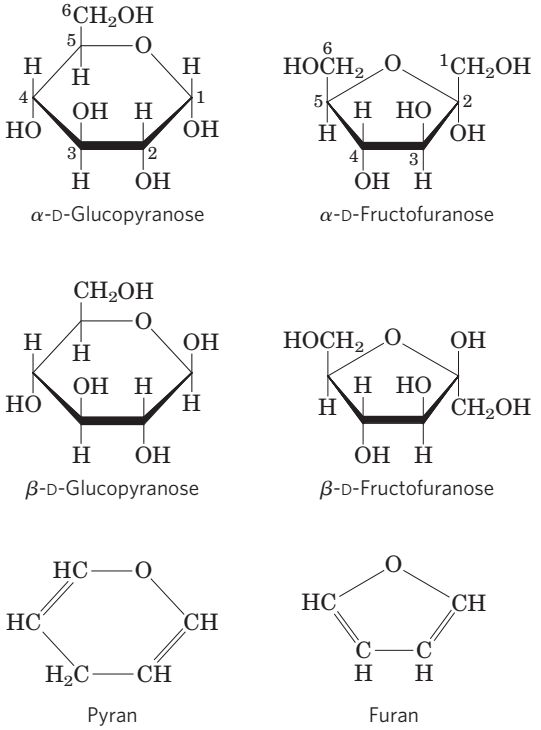
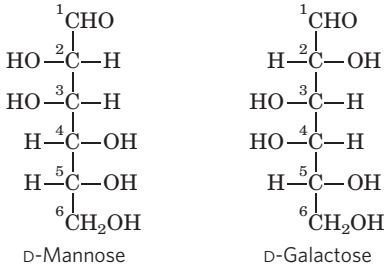


FIGURE 7–7 Pyranoses and furanoses. The pyranose forms of D-glucose and the furanose forms of D-fructose are shown here as Haworth perspective formulas. The edges of the ring nearest the reader are represented by bold lines. Hydroxyl groups below the plane of the ring in these Haworth perspectives would appear at the right side of a Fischer projection (compare with Fig. 7–6). Pyran and furan are shown for comparison.

WORKED EXAMPLE 7–1 Conversion of Fischer Projection to Haworth Perspective Formulas

Draw the Haworth perspective formulas for D-mannose and D-galactose.



Solution: Pyranoses are six-membered rings, so start with six-membered Haworth structures with the oxygen atom at the top right. Number the carbon atoms clockwise, starting with the aldose carbon. For mannose, place the hydroxyls on C-2, C-3, and C-4 above, above, and below the ring, respectively (because in the Fischer projection they are on the left, left, and right sides of the mannose structure). For D-galactose, the hydroxyls are oriented below, above, and above for C-2, C-3, and C-4, respectively. The hydroxyl at C-1 can be either up or down; there are two possible configurations, α and β, at this carbon.

WORKED EXAMPLE 7–2 Drawing Haworth Perspective Formulas of Sugar Isomers

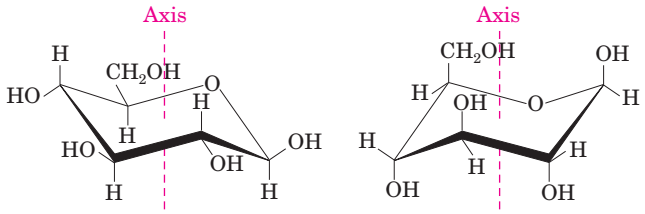
Draw the Haworth perspective formulas for α-D-mannose and β-L-galactose.

Solution: The Haworth perspective formula of D-mannose from Worked Example 7–1 can have the hydroxyl group at C-1 pointing either up or down. According to the Key Convention, for the α form, the C-1 hydroxyl is pointing down when C-6 is up, as it is in D-mannose.

For β-L-galactose, use the Fischer representation of D-galactose (see Worked Example 7–1) to draw the correct Fischer representation of L-galactose, which is its mirror image: the hydroxyls at C-2, C-3, C-4, and C-5 are on the left, right, right, and left sides, respectively. Now draw the Haworth perspective, a six-membered ring in which the —OH groups on C-2, C-3, and C-4 are oriented up, down, and down, respectively, because in the Fischer representation they are on the left, right, and right sides. Because it is the β form, the —OH on the anomeric carbon points down (same side as C-5).

The α and β anomers of D-glucose interconvert in aqueous solution by a process called **mutarotation**, in which one ring form (say, the α anomer) opens briefly into the linear form, then closes again to produce the β anomer (Fig. 7–6). Thus, a solution of β-D-glucose and a solution of α-D-glucose eventually form identical equilibrium mixtures having identical optical properties. This mixture consists of about one-third α-D-glucose, two-thirds β-D-glucose, and very small amounts of the linear and five-membered ring (glucofuranose) forms.

Haworth perspective formulas like those in Figure 7–7 are commonly used to show the stereochemistry of ring forms of monosaccharides. However, the six-membered pyranose ring is not planar, as Haworth perspectives suggest, but tends to assume either of two "chair" conformations **(Fig. 7–8)**. Recall from Chapter 1 (pp. 18–19) that two *conformations* of a molecule are interconvertible without the breakage of covalent bonds, whereas two *configurations* can be interconverted only by breaking a covalent bond. To interconvert α and β configurations, the bond involving the ring oxygen atom would have to be broken, but interconversion of the two chair forms (which are *conformers*) does not require bond breakage and does not change configurations at any of the ring carbons. The specific three-dimensional structures of the monosaccharide units are important in determining the biological properties and functions of some polysaccharides, as we shall see.



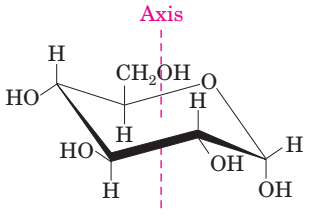


FIGURE 7–8 Conformational formulas of pyranoses. (a) Two chair forms of the pyranose ring of β-D-glucopyranose. Two *conformers* such as these are not readily interconvertible; an input of about 46 kJ of energy per mole of sugar is required to force the interconversion of chair forms. Another conformation, the "boat" (not shown), is seen only in derivatives with very bulky substituents. **(b)** The preferred chair conformation of α-D-glucopyranose.

Organisms Contain a Variety of Hexose Derivatives

In addition to simple hexoses such as glucose, galactose, and mannose, there are a number of sugar derivatives in which a hydroxyl group in the parent compound is replaced with another substituent, or a carbon atom is oxidized to a carboxyl group **(Fig. 7–9)**. In glucosamine, galactosamine, and mannosamine, the hydroxyl at C-2 of the parent compound is replaced with an amino group. The amino group is commonly condensed with acetic

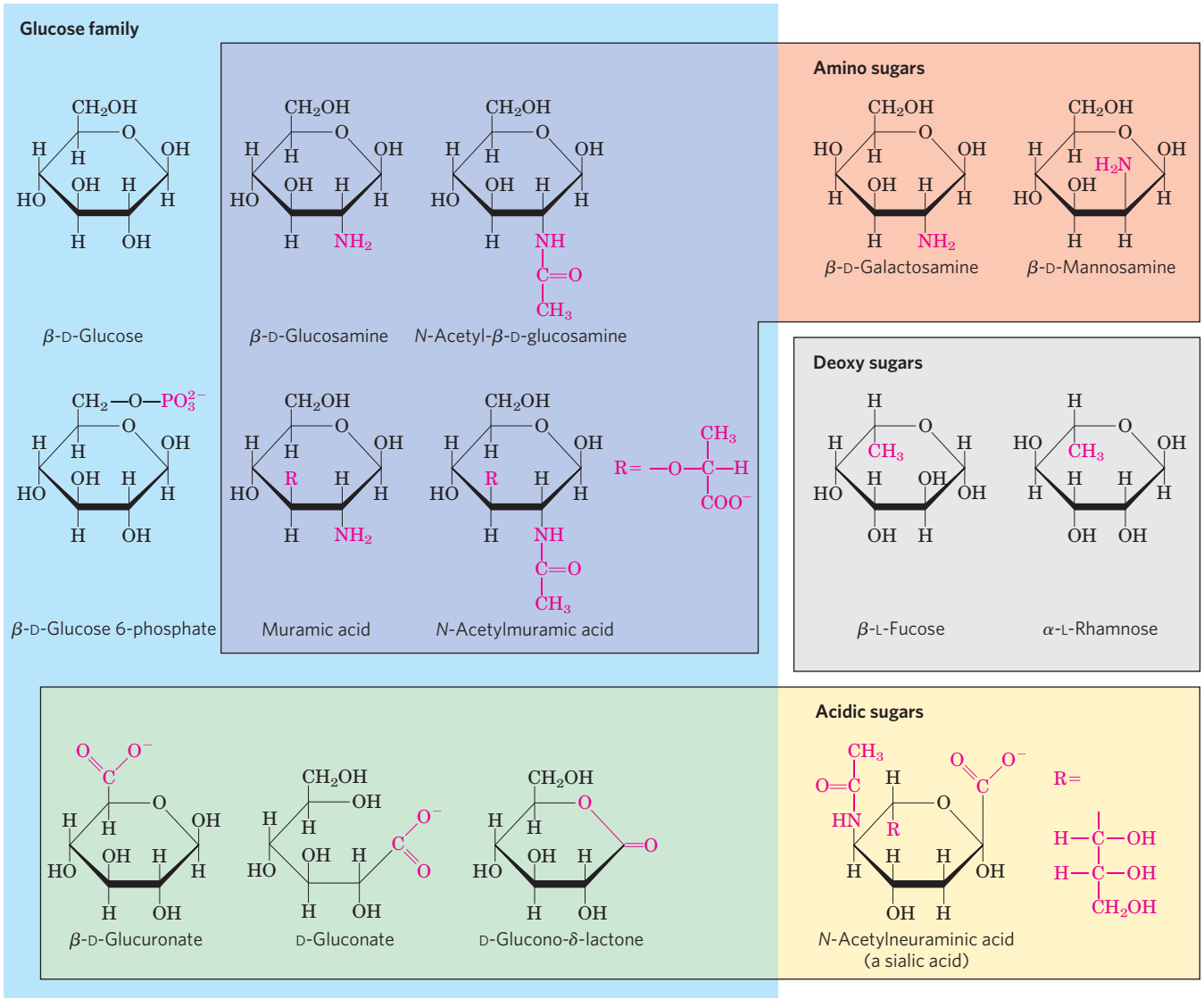


FIGURE 7–9 Some hexose derivatives important in biology. In amino sugars, an $—NH_2$ group replaces one of the —OH groups in the parent hexose. Substitution of —H for —OH produces a deoxy sugar; note that the deoxy sugars shown here occur in nature as the L isomers. The acidic sugars contain a carboxylate group, which confers a negative charge at neutral pH. D-Glucono-δ-lactone results from formation of an ester linkage between the C-1 carboxylate group and the C-5 (also known as the δ carbon) hydroxyl group of D-gluconate.

acid, as in *N*-acetylglucosamine. This glucosamine derivative is part of many structural polymers, including those of the bacterial cell wall. The substitution of a hydrogen for the hydroxyl group at C-6 of L-galactose or L-mannose produces L-fucose or L-rhamnose, respectively. L-Fucose is found in the complex oligosaccharide components of glycoproteins and glycolipids; L-rhamnose is found in plant polysaccharides.

Oxidation of the carbonyl (aldehyde) carbon of glucose to the carboxyl level produces gluconic acid, used in medicine as an innocuous counterion with which to administer positively charged drugs (such as quinine) or ions (such as Ca^{2+}). Other aldoses yield other **aldonic acids**. Oxidation of the carbon at the other end of the carbon chain—C-6 of glucose, galactose, or mannose—forms the corresponding **uronic acid**: glucuronic, galacturonic, or mannuronic acid. Both aldonic and uronic acids form stable intramolecular esters called lactones (Fig. 7–9, lower left). The sialic acids are a family of sugars with the same nine-carbon backbone. One of them, *N*-acetylneuraminic acid (often referred to simply as "sialic acid"), is a derivative of *N*-acetylmannosamine that occurs in many glycoproteins and glycolipids on animal cell surfaces, providing sites of recognition by other cells or extracellular carbohydrate-binding proteins. The carboxylic acid groups of the acidic sugar derivatives are ionized at pH 7, and the compounds are therefore correctly named as the carboxylates—glucuronate, galacturonate, and so forth.

In the synthesis and metabolism of carbohydrates, the intermediates are very often not the sugars

BOX 7–1 MEDICINE Blood Glucose Measurements in the Diagnosis and Treatment of Diabetes

Glucose is the principal fuel for the brain. When the amount of glucose reaching the brain is too low, the consequences can be dire: lethargy, coma, permanent brain damage, and death (see Fig. 23–24). Animals have evolved complex hormonal mechanisms to ensure that the concentration of glucose in the blood remains high enough (about 5 mM) to satisfy the brain's needs, but not too high, because elevated blood glucose can also have serious physiological consequences.

Individuals with insulin-dependent diabetes mellitus do not produce sufficient insulin, the hormone that normally serves to reduce blood glucose concentration, and if the diabetes is untreated their blood glucose levels may rise to severalfold higher than normal. These high glucose levels are believed to be at least one cause of the serious long-term consequences of untreated diabetes—kidney failure, cardiovascular disease, blindness, and impaired wound healing—so one goal of therapy is to provide just enough insulin (by injection) to keep blood glucose levels near normal. To maintain the correct balance of exercise, diet, and insulin for the individual, blood glucose concentration needs to be measured several times a day, and the amount of insulin injected adjusted appropriately.

The concentrations of glucose in blood and urine can be determined by a simple assay for reducing sugar, such as Fehling's reaction, which for many years was used as a diagnostic test for diabetes. Modern measurements require just a drop of blood, added to a test strip containing the enzyme glucose oxidase, which catalyzes the following reaction:

$$\text{D-Glucose} + O_2 \xrightarrow{\text{glucose oxidase}} \text{D-Glucono-}\delta\text{-lactone} + H_2O_2$$

A second enzyme, a peroxidase, catalyzes the reaction of the H_2O_2 with colorless compound to create a colored product, which is quantified with a simple photometer that reads out the blood glucose concentration.

Because blood glucose levels change with the timing of meals and exercise, single-time measurements do not reflect the *average* blood glucose over hours and days, so dangerous increases may go undetected. The average glucose concentration can be assessed by looking at its effect on hemoglobin, the oxygen-carrying protein in erythrocytes (p. 163). Transporters in the erythrocyte membrane equilibrate intracellular and plasma glucose concentrations, so hemoglobin is constantly exposed to glucose at whatever concentration is present in the blood. A nonenzymatic reaction occurs between glucose and primary amino groups in hemoglobin (either the amino-terminal Val or the ε-amino groups of Lys residues) (Fig. 1). The rate of this process is proportional to the concentration of glucose, so the reaction can be used as the basis for estimating the average blood glucose level over weeks. The amount of glycated hemoglobin (GHB) present at any time reflects the average blood glucose concentration over the circulating "lifetime" of the erythrocyte (about 120 days), although the concentration in the last two weeks is the most important in setting the level of GHB.

The extent of **hemoglobin glycation** (so named to distinguish it from glycosylation, the *enzymatic* transfer of glucose to a protein) is measured clinically by extracting hemoglobin from a small sample of blood and separating GHB from unmodified hemoglobin electrophoretically, taking advantage of the charge difference resulting from modification of the amino group(s). Normal GHB values are about 5% of total hemoglobin (corresponding to blood glucose of

themselves but their phosphorylated derivatives. Condensation of phosphoric acid with one of the hydroxyl groups of a sugar forms a phosphate ester, as in glucose 6-phosphate (Fig. 7–9), the first metabolite in the pathway by which most organisms oxidize glucose for energy. Sugar phosphates are relatively stable at neutral pH and bear a negative charge. One effect of sugar phosphorylation within cells is to trap the sugar inside the cell; most cells do not have plasma membrane transporters for phosphorylated sugars. Phosphorylation also activates sugars for subsequent chemical transformation. Several important phosphorylated derivatives of sugars are components of nucleotides (discussed in the next chapter).

Monosaccharides Are Reducing Agents

Monosaccharides can be oxidized by relatively mild oxidizing agents such as cupric (Cu^{2+}) ion. The carbonyl carbon is oxidized to a carboxyl group. Glucose and other sugars capable of reducing cupric ion are called **reducing sugars**. Cupric ion oxidizes glucose and certain other sugars to a complex mixture of carboxylic acids. This is the basis of Fehling's reaction, a semi-quantitative test for the presence of reducing sugar that for many years was used to detect and measure elevated glucose levels in people with diabetes mellitus. Today, more sensitive methods that involve an immobilized enzyme on a test strip are used; they require only a single drop of blood (Box 7–1). ■

120 mg/100 mL). In people with untreated diabetes, however, this value may be as high as 13%, indicating an average blood glucose level of about 300 mg/100 mL—dangerously high. One criterion for success in an individual program of insulin therapy (the timing, frequency, and amount of insulin injected) is maintaining GHB values at about 7%.

In the hemoglobin glycation reaction, the first step (formation of a Schiff base) is followed by a series of rearrangements, oxidations, and dehydrations of the carbohydrate moiety to produce a heterogeneous mixture of AGEs, *a*dvanced *g*lycation *e*nd products. These products can leave the erythrocyte and form covalent cross-links between proteins, interfering with normal protein function (Fig. 1). The accumulation of relatively high concentrations of AGEs in people with diabetes may, by cross-linking critical proteins, cause the damage to the kidneys, retinas, and cardiovascular system that characterizes the disease. This pathogenic process is a potential target for drug action.

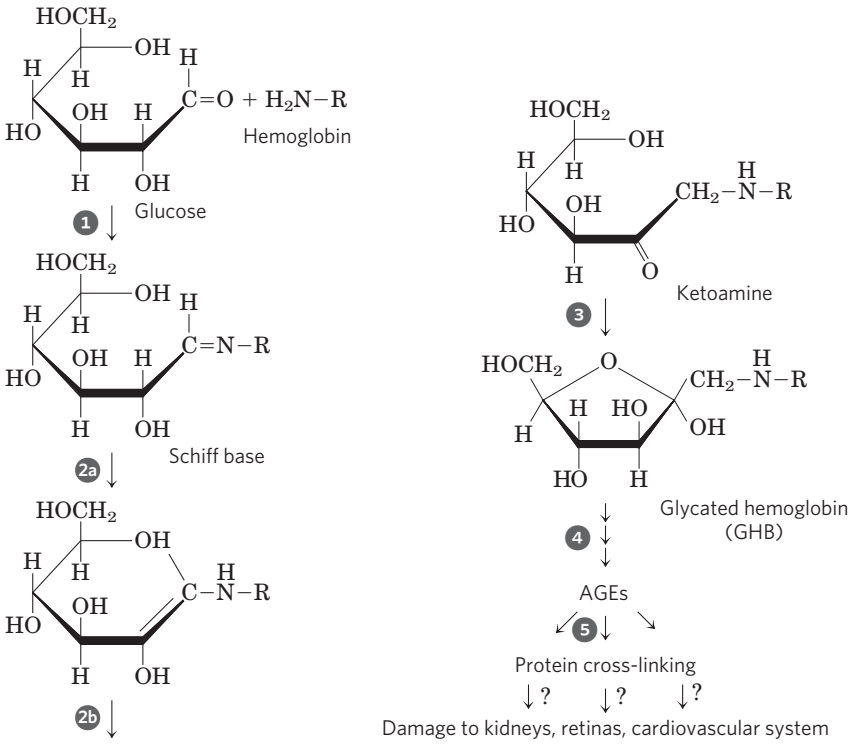


FIGURE 1 The nonenzymatic reaction of glucose with a primary amino group in hemoglobin begins with ❶ formation of a Schiff base, which ❷ undergoes a rearrangement to generate a stable product; ❸ this ketoamine can further cyclize to yield GHB. ❹ Subsequent reactions generate advanced glycation end products (AGEs), such as ε-*N*-carboxymethyllysine and methylglyoxal, compounds that ❺ can damage other proteins by cross-linking them, causing pathological changes.

Disaccharides Contain a Glycosidic Bond

Disaccharides (such as maltose, lactose, and sucrose) consist of two monosaccharides joined covalently by an ***O*-glycosidic bond**, which is formed when a hydroxyl group of one sugar molecule, typically cyclic, reacts with the anomeric carbon of the other **(Fig. 7–10)**. This reaction represents the formation of an acetal from a hemiacetal (such as glucopyranose) and an alcohol (a hydroxyl group of the second sugar molecule) (Fig. 7–5), and the resulting compound is called a glycoside. Glycosidic bonds are readily hydrolyzed by acid but resist cleavage by base. Thus disaccharides can be hydrolyzed to yield their free monosaccharide components by boiling with dilute acid. ***N*-glycosyl bonds** join the anomeric carbon of a sugar to a nitrogen atom in glycoproteins (see Fig. 7–30) and nucleotides (see Fig. 8–1).

The oxidation of a sugar by cupric ion (the reaction that defines a reducing sugar) occurs only with the linear form, which exists in equilibrium with the cyclic form(s). When the anomeric carbon is involved in a glycosidic bond (that is, when the compound is a full acetal or ketal; see Fig. 7–5), the easy interconversion of linear and cyclic forms shown in Figure 7–6 is prevented. Because the carbonyl carbon can be oxidized only when the sugar is in its linear form, formation of a glycosidic bond renders a sugar nonreducing. In describing disaccharides or polysaccharides, the end of a chain with a free anomeric carbon (one not involved in a glycosidic bond) is commonly called the **reducing end**.

The disaccharide maltose (Fig. 7–10) contains two D-glucose residues joined by a glycosidic linkage between C-1 (the anomeric carbon) of one glucose residue and C-4 of the other. Because the disaccharide retains a free anomeric carbon (C-1 of the glucose residue on the right in Fig. 7–10), maltose is a reducing sugar. The configuration of the anomeric carbon atom in the glycosidic linkage is α. The glucose residue with the free anomeric carbon is capable of existing in α- and β-pyranose forms.

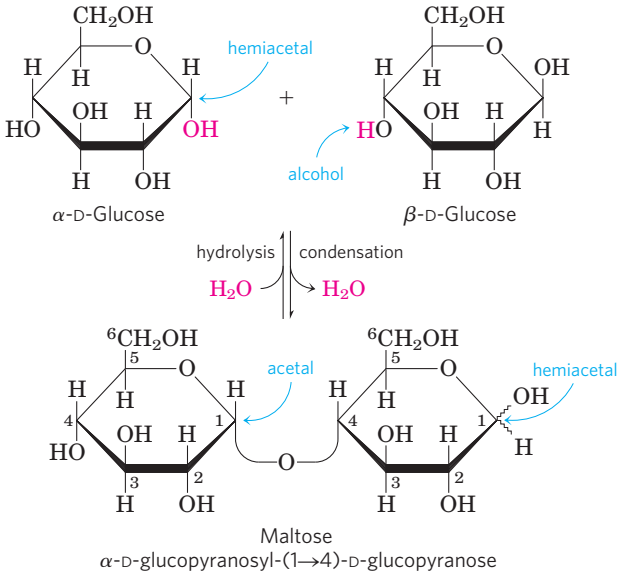


FIGURE 7–10 Formation of maltose. A disaccharide is formed from two monosaccharides (here, two molecules of D-glucose) when an —OH (alcohol) of one monosaccharide molecule (right) condenses with the intramolecular hemiacetal of the other (left), with elimination of H_2O and formation of a glycosidic bond. The reversal of this reaction is hydrolysis—attack by H_2O on the glycosidic bond. The maltose molecule shown here retains a reducing hemiacetal at the C-1 not involved in the glycosidic bond. Because mutarotation interconverts the α and β forms of the hemiacetal, the bonds at this position are sometimes depicted with wavy lines, as shown here, to indicate that the structure may be either α or β.

KEY CONVENTION: To name reducing disaccharides such as maltose unambiguously, and especially to name more complex oligosaccharides, several rules are followed. By convention, the name describes the compound written with its nonreducing end to the left, and we can "build up" the name in the following order. (1) Give the configuration (α or β) at the anomeric carbon joining the first monosaccharide unit (on the left) to the second. (2) Name the nonreducing residue; to distinguish five- and six-membered ring structures, insert "furano" or "pyrano" into the name. (3) Indicate in parentheses the two carbon atoms joined by the glycosidic bond, with an arrow connecting the two numbers; for example, (1→4) shows that C-1 of the first-named sugar residue is joined to C-4 of the second. (4) Name the second residue. If there is a third residue, describe the second glycosidic bond by the same conventions. (To shorten the description of complex polysaccharides, three-letter abbreviations or colored symbols for the monosaccharides are often used, as given in Table 7–1.) Following this convention for naming oligosaccharides, maltose is α-D-glucopyranosyl-(1→4)-D-glucopyranose. Because

TABLE 7–1 Symbols and Abbreviations for Common Monosaccharides and Some of Their Derivatives

Abequose	Abe	Glucuronic acid	◆ GlcA
Arabinose	Ara	Galactosamine	□ GalN
Fructose	Fru	Glucosamine	◥ GlcN
Fucose	▲ Fuc	*N*-Acetylgalactosamine	▨ GalNAc
Galactose	○ Gal	*N*-Acetylglucosamine	■ GlcNAc
Glucose	● Glc	Iduronic acid	◇ IdoA
Mannose	● Man	Muramic acid	Mur
Rhamnose	Rha	*N*-Acetylmuramic acid	Mur2Ac
Ribose	Rib	*N*-Acetylneuraminic acid (a sialic acid)	◆ Neu5Ac
Xylose	★ Xyl		

Note: In a commonly used convention, hexoses are represented as circles, *N*-acetylhexosamines as squares, and hexosamines as squares divided diagonally. All sugars with the "gluco" configuration are blue, those with the "galacto" configuration are yellow, and "manno" sugars are green. Other substituents can be added as needed: sulfate (S), phosphate (P), *O*-acetyl (OAc), or *O*-methyl (OMe).

most sugars encountered in this book are the D enantiomers and the pyranose form of hexoses predominates, we generally use a shortened version of the formal name of such compounds, giving the configuration of the anomeric carbon and naming the carbons joined by the glycosidic bond. In this abbreviated nomenclature, maltose is Glc(α1→4)Glc. ■

The disaccharide lactose **(Fig. 7–11)**, which yields D-galactose and D-glucose on hydrolysis, occurs naturally in milk. The anomeric carbon of the glucose residue is available for oxidation, and thus lactose is a reducing disaccharide. Its abbreviated name is Gal(β1→4)Glc. Sucrose (table sugar) is a disaccharide of glucose and fructose. It is formed by plants but not by animals. In contrast to maltose and lactose, sucrose contains no free anomeric carbon atom; the anomeric carbons of both monosaccharide units are involved in the glycosidic bond (Fig. 7–11). Sucrose is therefore a nonreducing sugar, and its stability toward oxidation makes it a suitable molecule for the storage and transport of energy in plants. In the abbreviated nomenclature, a double-headed arrow connects the symbols specifying the anomeric carbons and their configurations. For example, the abbreviated name of sucrose is either Glc(α1↔2β)Fru or Fru(β2↔1α)Glc. Sucrose is a major intermediate product of photosynthesis; in many plants it is the principal form in which sugar is transported from the leaves to other parts of the plant body. Trehalose, Glc(α1↔1α)Glc (Fig. 7–11)—a disaccharide of D-glucose that, like sucrose, is a nonreducing sugar—is a major constituent of the circulating fluid (hemolymph) of insects, serving as an energy-storage compound. Lactose gives milk its sweetness, and sucrose, of course, is table sugar. Trehalose is also used commercially as a sweetener. Box 7–2 explains how humans detect sweetness, and how artificial sweeteners such as aspartame act.

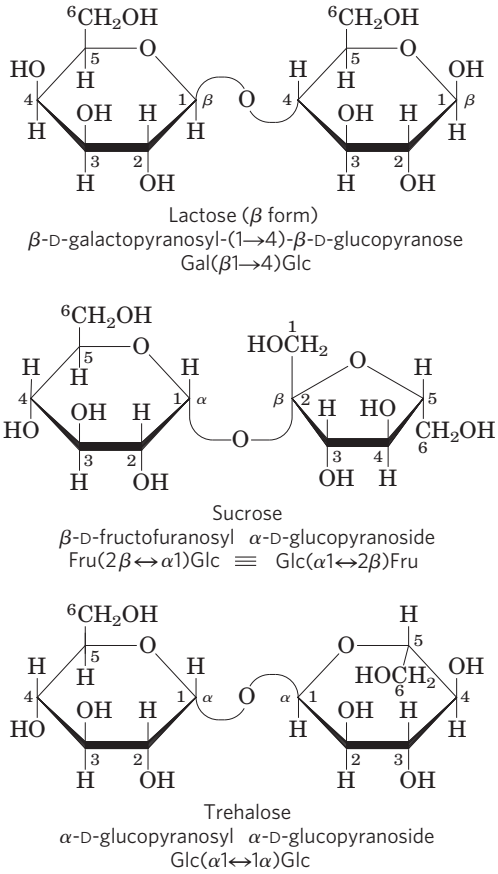


FIGURE 7–11 Two common disaccharides. Like maltose in Figure 7–10, these are shown as Haworth perspectives. The common name, full systematic name, and abbreviation are given for each disaccharide. Formal nomenclature for sucrose names glucose as the parent glycoside, although it is typically depicted as shown, with glucose on the left. The two abbreviated symbols shown for sucrose are equivalent (≡).

SUMMARY 7.1 Monosaccharides and Disaccharides

- Sugars (also called saccharides) are compounds containing an aldehyde or ketone group and two or more hydroxyl groups.
- Monosaccharides generally contain several chiral carbons and therefore exist in a variety of stereochemical forms, which may be represented on paper as Fischer projections. Epimers are sugars that differ in configuration at only one carbon atom.
- Monosaccharides commonly form internal hemiacetals or hemiketals, in which the aldehyde or ketone group joins with a hydroxyl group of the same molecule, creating a cyclic structure; this can be represented as a Haworth perspective formula. The carbon atom originally found in the aldehyde or ketone group (the anomeric carbon) can assume either of two configurations, α and β, which are interconvertible by mutarotation. In the linear form of the monosaccharide, which is in equilibrium with the cyclic forms, the anomeric carbon is easily oxidized, making the compound a reducing sugar.
- A hydroxyl group of one monosaccharide can add to the anomeric carbon of a second monosaccharide to form an acetal called a glycoside. In this disaccharide, the glycosidic bond protects the anomeric carbon from oxidation, making it a nonreducing sugar.
- Oligosaccharides are short polymers of several monosaccharides joined by glycosidic bonds. At one end of the chain, the reducing end, is a monosaccharide unit with its anomeric carbon not involved in a glycosidic bond.
- The common nomenclature for di- or oligosaccharides specifies the order of monosaccharide units, the configuration at each anomeric carbon, and the carbon atoms involved in the glycosidic linkage(s).

BOX 7–2 Sugar Is Sweet, and So Are . . . a Few Other Things

Sweetness is one of the five basic flavors that humans can taste (Fig. 1); the others are sour, bitter, salty, and umami. Sweet taste is detected by protein receptors in the plasma membranes of gustatory cells in the taste buds on the surface of the tongue. In humans, two closely related genes (*T1R2* and *T1R3*) encode sweetness receptors (Fig. 2). When a molecule with a compatible structure binds these receptors on a gustatory cell's extracellular domain, it triggers a series of events in the cell (including activation of a GTP-binding protein; see Fig. 12–42) that lead to an electrical signal being sent to the brain that is interpreted there as "sweet." During evolution, there has probably been selection for the ability to taste compounds found in foods containing important nutrients, such as the carbohydrates that are major fuels for most organisms. Most simple sugars, including sucrose, glucose, and fructose,

FIGURE 1 A strong stimulus for the sweetness receptors.

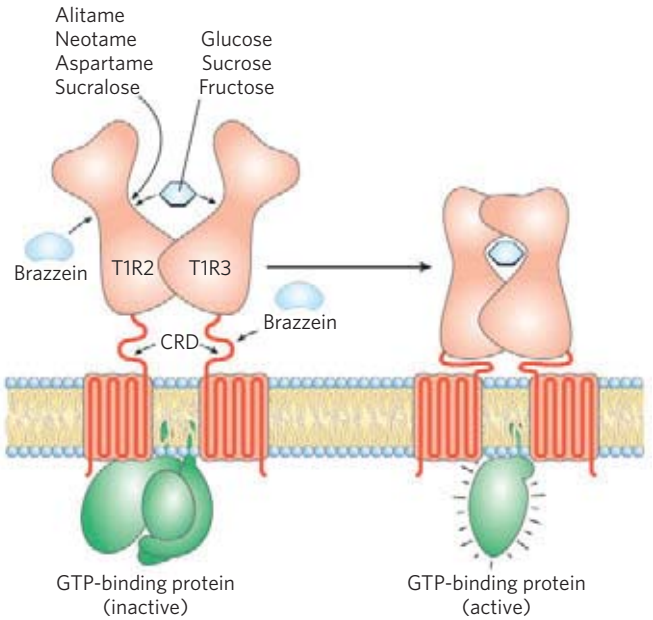


FIGURE 2 The receptor for sweet-tasting substances, with its regions of interaction (short arrows) with various sweet-tasting compounds indicated. Each receptor has an extracellular domain, a cysteine-rich domain (CRD), and a membrane domain with seven transmembrane helices, a common feature of signaling receptors. Artificial sweeteners bind to only one of the two receptor subunits; natural sugars bind to both. See Chapter 1, Problem 14, for the structures of many of these artificial sweeteners.

7.2 Polysaccharides

Most carbohydrates found in nature occur as polysaccharides, polymers of medium to high molecular weight (M_r >20,000). Polysaccharides, also called **glycans**, differ from each other in the identity of their recurring monosaccharide units, in the length of their chains, in the types of bonds linking the units, and in the degree of branching. **Homopolysaccharides** contain only a single monomeric species; **heteropolysaccharides** contain two or more different kinds **(Fig. 7–12)**. Some homopolysaccharides serve as storage forms of monosaccharides that are used as fuels; starch and glycogen are homopolysaccharides of this type. Other homopolysaccharides (cellulose and chitin, for example) serve as structural elements in plant cell walls and animal exoskeletons. Heteropolysaccharides provide extracellular support for organisms of all kingdoms. For example, the rigid layer of the bacterial cell envelope (the peptidoglycan) is composed in part of a heteropolysaccharide built from two alternating monosaccharide units (see Fig. 20–30). In animal tissues, the extracellular space is

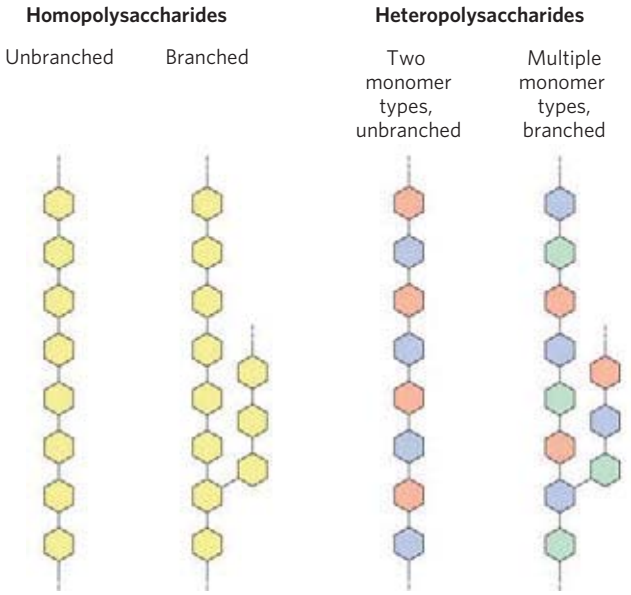


FIGURE 7–12 Homo- and heteropolysaccharides. Polysaccharides may be composed of one, two, or several different monosaccharides, in straight or branched chains of varying length.

taste sweet, but there are other classes of compounds that also bind the sweet receptors: the amino acids glycine, alanine, and serine are mildly sweet and harmless; nitrobenzene and ethylene glycol have a strong sweet taste, but are toxic. (See Box 18–2 for a remarkable medical mystery involving ethylene glycol poisoning.) Several natural products are extraordinarily sweet: stevioside, a sugar derivative isolated from the leaves of the stevia plant (*Stevia rebaudiana* Berton), is several hundred times sweeter than an equivalent amount of sucrose (table sugar), and the small (54 amino acids) protein brazzein, isolated from berries of the Oubli vine (*Pentadiplandra brazzeana* Baillon) in Gabon and Cameroon, is 17,000 times as sweet as sucrose on a molar basis. Presumably the sweet taste of the berries encourages their consumption by animals, which then disperse the seeds geographically so new plants may be established.

There is great interest in the development of artificial sweeteners as weight-reduction aids—compounds that give foods a sweet taste without adding the calories found in sugars. The artificial sweetener aspartame demonstrates the importance of stereochemistry in biology (Fig. 3). According to one simple model of sweetness receptor binding, binding involves three sites on the receptor: AH^+, B^-, and X. Site AH^+ contains some group (an alcohol or amine) that can form a hydrogen bond with a partial negative charge, such as a carbonyl oxygen, on the sweetener molecule; the carboxylic acid of aspartame contains such an oxygen. Site B^- contains a group with a partially negative oxygen available to hydrogen-bond with some partially positive atom on the sweetener molecule, such as the amine group of aspartame. Site X is oriented perpendicular to the other two groups and is capable of interacting with a hydrophobic patch on the sweetener molecule, such as the benzene ring of aspartame.

When the steric match is correct, as on the left in Figure 3, the sweet receptor is stimulated and the signal "sweet" is conducted to the brain. When the match is not correct, as on the right in Figure 3, the sweet receptor is not stimulated; in fact, in this case, another receptor (for bitterness) is stimulated by the "wrong" stereoisomer of aspartame. Stereoisomerism really matters!

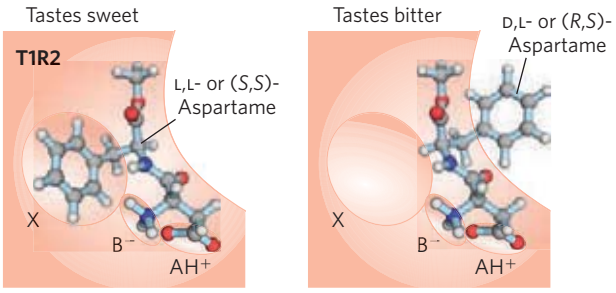


FIGURE 3 Stereochemical basis for the taste of two isomers of aspartame.

occupied by several types of heteropolysaccharides, which form a matrix that holds individual cells together and provides protection, shape, and support to cells, tissues, and organs.

Unlike proteins, polysaccharides generally do not have defining molecular weights. This difference is a consequence of the mechanisms of assembly of the two types of polymer. As we shall see in Chapter 27, proteins are synthesized on a template (messenger RNA) of defined sequence and length, by enzymes that follow the template exactly. For polysaccharide synthesis there is no template; rather, the program for polysaccharide synthesis is intrinsic to the enzymes that catalyze the polymerization of the monomeric units, and there is no specific stopping point in the synthetic process; the products thus vary in length.

Some Homopolysaccharides Are Stored Forms of Fuel

The most important storage polysaccharides are starch in plant cells and glycogen in animal cells. Both polysaccharides occur intracellularly as large clusters or granules. Starch and glycogen molecules are heavily hydrated, because they have many exposed hydroxyl groups available to hydrogen-bond with water. Most plant cells have the ability to form starch (see Fig. 20–2), and starch storage is especially abundant in tubers (underground stems), such as potatoes, and in seeds.

Starch contains two types of glucose polymer, amylose and amylopectin **(Fig. 7–13)**. Amylose consists of long, unbranched chains of D-glucose residues connected by ($\alpha 1 \rightarrow 4$) linkages (as in maltose). Such chains vary in molecular weight from a few thousand to more than a million. Amylopectin also has a high molecular weight (up to 200 million) but unlike amylose is highly branched. The glycosidic linkages joining successive glucose residues in amylopectin chains are ($\alpha 1 \rightarrow 4$); the branch points (occurring every 24 to 30 residues) are ($\alpha 1 \rightarrow 6$) linkages.

Glycogen is the main storage polysaccharide of animal cells. Like amylopectin, glycogen is a polymer of ($\alpha 1 \rightarrow 4$)-linked subunits of glucose, with ($\alpha 1 \rightarrow 6$)-linked branches, but glycogen is more extensively branched (on average, every 8 to 12 residues) and more compact than

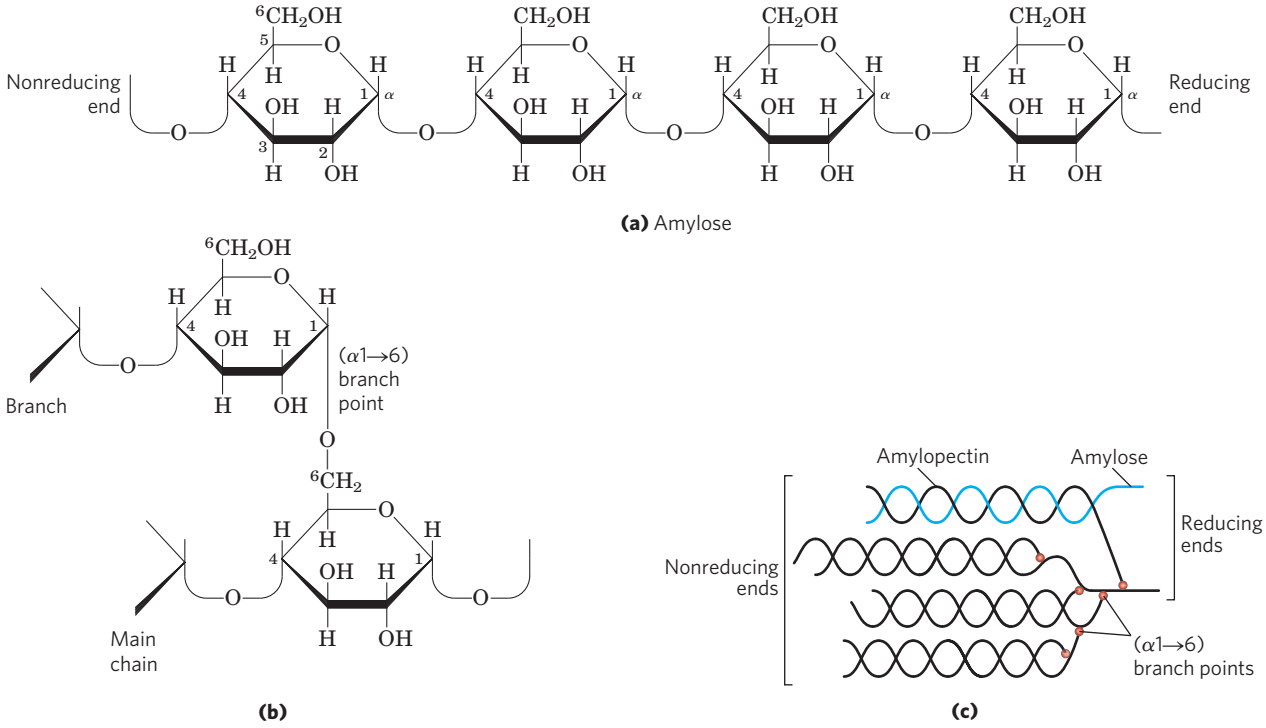


FIGURE 7–13 Glycogen and starch. (a) A short segment of amylose, a linear polymer of D-glucose residues in (α1⟶4) linkage. A single chain can contain several thousand glucose residues. Amylopectin has stretches of similarly linked residues between branch points. Glycogen has the same basic structure, but has more branching than amylopectin. **(b)** An (α1⟶6) branch point of glycogen or amylopectin. **(c)** A cluster of amylose and amylopectin like that believed to occur in starch granules. Strands of amylopectin (black) form double-helical structures with each other or with amylose strands (blue). Amylopectin has frequent (α1⟶6) branch points (red). Glucose residues at the nonreducing ends of the outer branches are removed enzymatically during the mobilization of starch for energy production. Glycogen has a similar structure but is more highly branched and more compact.

starch. Glycogen is especially abundant in the liver, where it may constitute as much as 7% of the wet weight; it is also present in skeletal muscle. In hepatocytes glycogen is found in large granules, which are themselves clusters of smaller granules composed of single, highly branched glycogen molecules with an average molecular weight of several million. Such glycogen granules also contain, in tightly bound form, the enzymes responsible for the synthesis and degradation of glycogen.

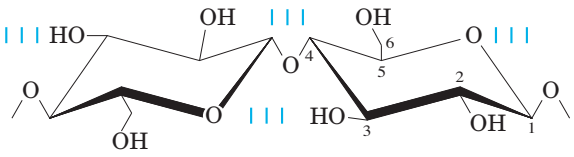
Because each branch in glycogen ends with a nonreducing sugar unit, a glycogen molecule with n branches has $n + 1$ nonreducing ends, but only one reducing end. When glycogen is used as an energy source, glucose units are removed one at a time from the nonreducing ends. Degradative enzymes that act only at nonreducing ends can work simultaneously on the many branches, speeding the conversion of the polymer to monosaccharides.

Why not store glucose in its monomeric form? It has been calculated that hepatocytes store glycogen equivalent to a glucose concentration of 0.4 M. The actual concentration of glycogen, which is insoluble and contributes little to the osmolarity of the cytosol, is about 0.01 μM. If the cytosol contained 0.4 M glucose, the osmolarity would be threateningly elevated, leading to osmotic entry of water that might rupture the cell (see Fig. 2–13). Furthermore, with an intracellular glucose concentration of 0.4 M and an external concentration of about 5 mM (the concentration in the blood of a mammal), the free-energy change for glucose uptake into cells against this very high concentration gradient would be prohibitively large.

Dextrans are bacterial and yeast polysaccharides made up of (α1⟶6)-linked poly-D-glucose; all have (α1⟶3) branches, and some also have (α1⟶2) or (α1⟶4) branches. Dental plaque, formed by bacteria growing on the surface of teeth, is rich in dextrans, which are adhesive and allow the bacteria to stick to teeth and to each other. Dextrans also provide a source of glucose for bacterial metabolism. Synthetic dextrans are used in several commercial products (for example, Sephadex) that serve in the fractionation of proteins by size-exclusion chromatography (see Fig. 3–17b). The dextrans in these products are chemically cross-linked to form insoluble materials of various sizes.

Some Homopolysaccharides Serve Structural Roles

Cellulose, a fibrous, tough, water-insoluble substance, is found in the cell walls of plants, particularly in stalks, stems, trunks, and all the woody portions of the plant body. Cellulose constitutes much of the mass of wood, and cotton is almost pure cellulose. Like amylose, the



(β1→4)-linked D-glucose units

FIGURE 7–14 Cellulose. Two units of a cellulose chain; the D-glucose residues are in (β1→4) linkage. The rigid chair structures can rotate relative to one another.

cellulose molecule is a linear, unbranched homopolysaccharide, consisting of 10,000 to 15,000 D-glucose units. But there is a very important difference: in cellulose the glucose residues have the β configuration **(Fig. 7–14)**, whereas in amylose the glucose is in the α configuration. The glucose residues in cellulose are linked by (β1→4) glycosidic bonds, in contrast to the (α1→4) bonds of amylose. This difference causes individual molecules of cellulose and amylose to fold differently in space, giving them very different macroscopic structures and physical properties (see below). The tough, fibrous nature of cellulose makes it useful in such commercial products as cardboard and insulation material, and it is a major constituent of cotton and linen fabrics. Cellulose is also the starting material for the commercial production of cellophane and rayon.

Glycogen and starch ingested in the diet are hydrolyzed by α-amylases and glycosidases, enzymes in saliva and the intestine that break (α1→4) glycosidic bonds between glucose units. Most vertebrate animals cannot use cellulose as a fuel source, because they lack an enzyme to hydrolyze the (β1→4) linkages. Termites readily digest cellulose (and therefore wood), but only because their intestinal tract harbors a symbiotic microorganism, *Trichonympha*, that secretes cellulase, which hydrolyzes the (β1→4) linkages **(Fig. 7–15)**. Molecular genetic studies have revealed that genes encoding cellulose-degrading enzymes are present in the genomes of a wide range of invertebrate animals, including arthropods and nematodes. There is one important exception to the absence of cellulase in vertebrates: ruminant animals such as cattle, sheep, and goats harbor symbiotic microorganisms in the rumen (the first of their four stomach compartments) that can hydrolyze cellulose, allowing the animal to degrade dietary cellulose from soft grasses, but not from woody plants. Fermentation in the rumen yields acetate, propionate, and β-hydroxybutyrate, which the animal uses to synthesize the sugars in milk (p. 560).

FIGURE 7–15 Cellulose breakdown by *Trichonympha*, a protist in the gut of a wood-eating termite. *Trichonympha* produces the enzyme cellulase, which breaks the (β1→4) glycosidic bonds in cellulose, making wood a source of metabolizable sugar (glucose) for the protist and the termite. A number of invertebrates can digest cellulose, but only a few vertebrates (the ruminants, such as cattle, sheep, and goats); the ruminants are able to use cellulose as food because the first of their four stomach compartments (rumen) teems with bacteria and protists that secrete cellulase.

Biomass (such as switch grass) that is rich in cellulose can be used as starting material for the fermentation of carbohydrates to ethanol, to be used as a gasoline additive. The annual production of biomass on Earth (accomplished primarily by photosynthetic organisms) is the energetic equivalent of nearly a trillion barrels of crude oil, when converted to ethanol by fermentation. Because of their potential use in biomass conversion to bioenergy, cellulose-degrading enzymes such as cellulase are under vigorous investigation. Supramolecular complexes called cellulosomes, found on the outside surface of the bacterium *Clostridium cellulolyticum*, include the catalytic subunit of cellulase, along with proteins that hold one or more cellulase molecules to the bacterial surface, and a subunit that binds cellulose and positions it in the catalytic site.

A major fraction of photosynthetic biomass is the woody portion of plants and trees, which consists of cellulose plus several other polymers derived from carbohydrates that are not easily digestible, either chemically or biologically. Lignins, for example, make up some 30% of the mass of wood. Synthesized from precursors that include phenylalanine and glucose, lignins are complex polymers with covalent cross-links to cellulose that complicate the digestion of cellulose by cellulase. If woody plants are to be used in the production of ethanol from biomass, better means of digesting wood components will need to be found.

Chitin is a linear homopolysaccharide composed of *N*-acetylglucosamine residues in (β1→4) linkage **(Fig. 7–16)**. The only chemical difference from cellulose is the replacement of the hydroxyl group at C-2 with an acetylated amino group. Chitin forms extended fibers similar to those of cellulose, and like cellulose cannot be digested by vertebrates. Chitin is the principal component of the hard exoskeletons of nearly a million species of arthropods—insects, lobsters, and crabs, for example—and is probably the second most abundant polysaccharide, next to cellulose, in nature; an estimated 1 billion tons of chitin are produced each year in the biosphere.

Steric Factors and Hydrogen Bonding Influence Homopolysaccharide Folding

The folding of polysaccharides in three dimensions follows the same principles as those governing polypeptide

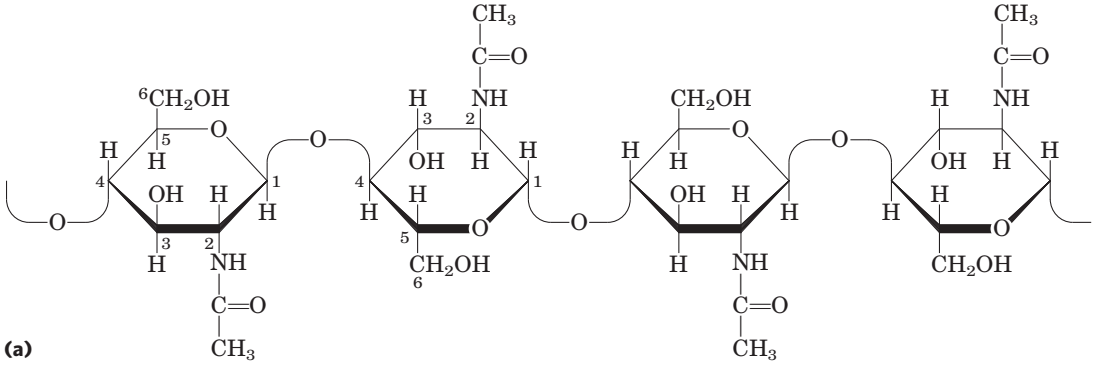


FIGURE 7–16 **Chitin.** **(a)** A short segment of chitin, a homopolymer of *N*-acetyl-D-glucosamine units in (β1⟶4) linkage. **(b)** A spotted June beetle (*Pelidnota punctata*), showing its surface armor (exoskeleton) of chitin.

structure: subunits with a more-or-less rigid structure dictated by covalent bonds form three-dimensional macromolecular structures that are stabilized by weak interactions within or between molecules, such as hydrogen bonds and hydrophobic and van der Waals interactions, and, for polymers with charged subunits, electrostatic interactions. Because polysaccharides have so many hydroxyl groups, hydrogen bonding has an especially important influence on their structure. Glycogen, starch, and cellulose are composed of pyranoside subunits (having six-membered rings), as are the oligosaccharides of glycoproteins and glycolipids, to be discussed later. Such molecules can be represented as a series of rigid pyranose rings connected by an oxygen atom bridging two carbon atoms (the glycosidic bond). There is, in principle, free rotation about both C—O bonds linking the residues (Fig. 7–14), but as in polypeptides (see Figs 4–2, 4–9), rotation about each bond is limited by steric hindrance by substituents. The three-dimensional structures of these molecules can be described in terms of the dihedral angles, ϕ and ψ, about the glycosidic bond **(Fig. 7–17)**, analogous to angles ϕ and ψ made by the peptide bond (see Fig. 4–2).

The bulkiness of the pyranose ring and its substituents, and electronic effects at the anomeric carbon, place constraints on the angles ϕ and ψ; thus certain conformations are much more stable than others, as can be shown on a map of energy as a function of ϕ and ψ **(Fig. 7–18)**.

The most stable three-dimensional structure for the (α1⟶4)-linked chains of starch and glycogen is a tightly coiled helix **(Fig. 7–19)**, stabilized by interchain hydrogen bonds. In amylose (with no branches) this structure is regular enough to allow crystallization and thus determination of the structure by x-ray diffraction. The average plane of each residue along the amylose chain forms a 60° angle with the average plane of the

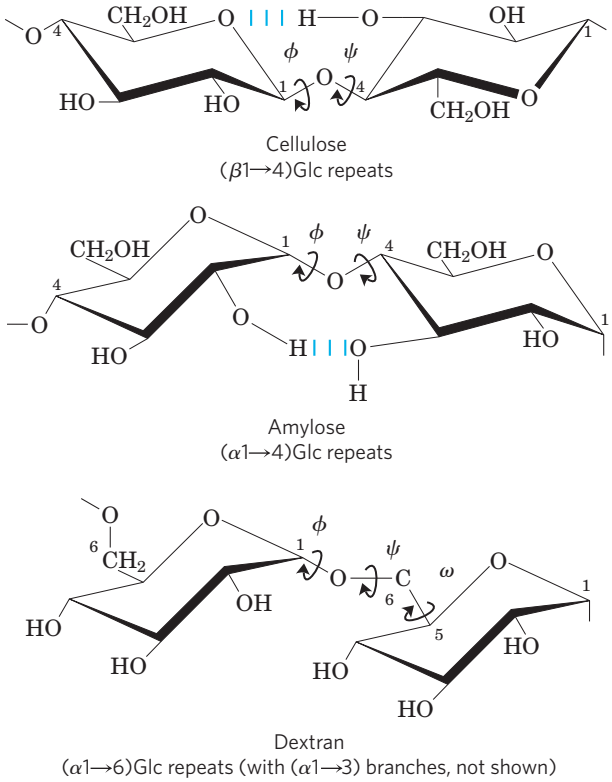
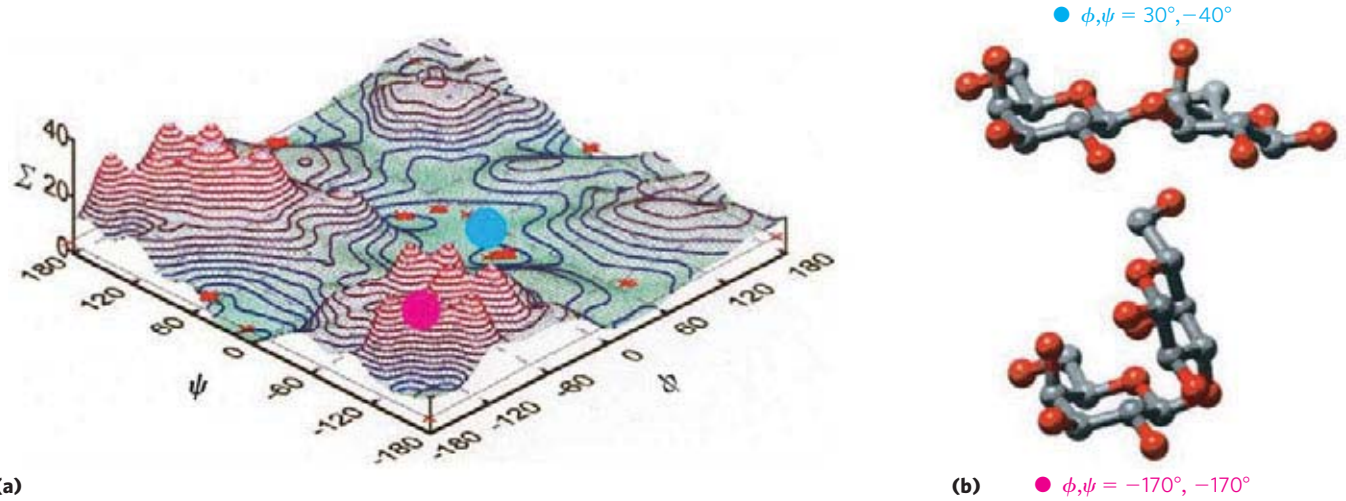


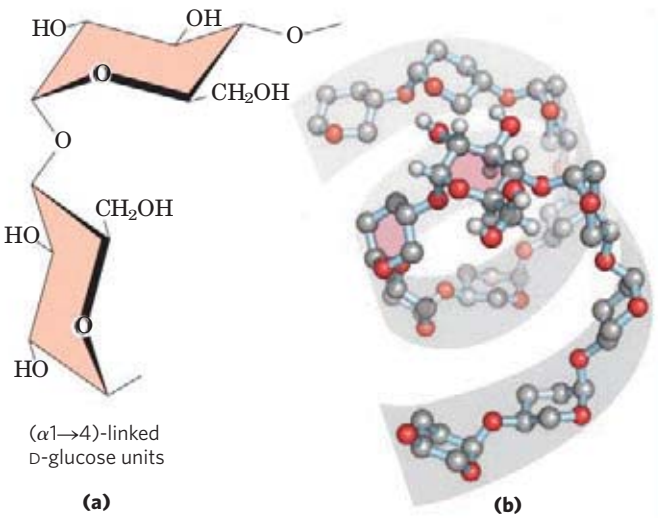
FIGURE 7–17 **Conformation at the glycosidic bonds of cellulose, amylose, and dextran.** The polymers are depicted as rigid pyranose rings joined by glycosidic bonds, with free rotation about these bonds. Note that in dextran there is also free rotation about the bond between C-5 and C-6 (torsion angle ω (omega)).



(a) (b)

FIGURE 7–18 A map of favored conformations for oligosaccharides and polysaccharides. The torsion angles ψ and ϕ (see Fig. 7–17), which define the spatial relationship between adjacent rings, can in principle have any value from 0° to 360°. In fact, some of the torsion angles would give conformations that are sterically hindered, whereas others give conformations that maximize hydrogen bonding. **(a)** When the relative energy (Σ) is plotted for each value of ϕ and ψ, with isoenergy ("same energy") contours drawn at intervals of 1 kcal/mol above the minimum energy state, the result is a map of preferred conformations. This is analogous to the Ramachandran plot for peptides (see Figs 4–3, 4–9). **(b)** Two energetic extremes for the disaccharide Gal(β1→3)Gal; these values fall on the energy diagram (a) as shown by the red and blue dots. The red dot indicates the least favored conformation; the blue dot, the most favored conformation. The known conformations of the three polysaccharides shown in Figure 7–17 have been determined by x-ray crystallography, and all fall within the lowest-energy regions of the map.

preceding residue, so the helical structure has six residues per turn. For amylose, the core of the helix is of precisely the right dimensions to accommodate iodine as complex ions (I_3^- and I_5^-), giving an intensely blue complex. This interaction is a common qualitative test for amylose.



(a) (b)

FIGURE 7–19 Helical structure of starch (amylose). (a) In the most stable conformation, with adjacent rigid chairs, the polysaccharide chain is curved, rather than linear as in cellulose (see Fig. 7–14). **(b)** A model of a segment of amylose; for clarity, the hydroxyl groups have been omitted from all but one of the glucose residues. Compare the two residues shaded in pink with the chemical structures in (a). The conformation of (α1→4) linkages in amylose, amylopectin, and glycogen causes these polymers to assume tightly coiled helical structures. These compact structures produce the dense granules of stored starch or glycogen seen in many cells (see Fig. 20–2).

For cellulose, the most stable conformation is that in which each chair is turned 180° relative to its neighbors, yielding a straight, extended chain. All —OH groups are available for hydrogen bonding with neighboring chains. With several chains lying side by side, a stabilizing network of interchain and intrachain hydrogen bonds produces straight, stable supramolecular fibers of great tensile strength **(Fig. 7–20)**. This property of cellulose has made it a useful substance to civilizations for millennia. Many manufactured products, including papyrus, paper, cardboard, rayon, insulating tiles, and a variety of other useful materials, are derived from cellulose. The water content of these materials is low because extensive interchain hydrogen bonding between cellulose molecules satisfies their capacity for hydrogen-bond formation.

Bacterial and Algal Cell Walls Contain Structural Heteropolysaccharides

The rigid component of bacterial cell walls (peptidoglycan) is a heteropolymer of alternating (β1→4)-linked *N*-acetylglucosamine and *N*-acetylmuramic acid residues (see Fig. 20–30). The linear polymers lie side by side in the cell wall, cross-linked by short peptides, the exact structure of which depends on the bacterial species. The peptide cross-links weld the polysaccharide chains into a strong sheath (peptidoglycan) that envelops the entire cell and prevents cellular swelling and lysis due to the osmotic entry of water. The enzyme

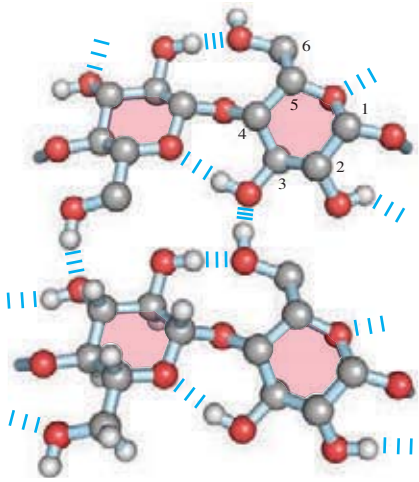
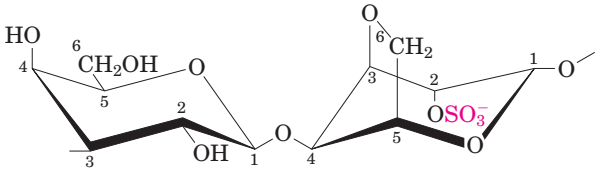


FIGURE 7–20 Cellulose chains. Scale drawing of segments of two parallel cellulose chains, showing the conformation of the D-glucose residues and the hydrogen-bond cross-links. In the hexose unit at the lower left, all hydrogen atoms are shown; in the other three hexose units, the hydrogens attached to carbon have been omitted for clarity, as they do not participate in hydrogen bonding.



Agarose
3)D-Gal(β1⟶4)3,6-anhydro-L-Gal2S (α1
repeating units

FIGURE 7–21 Agarose. The repeating unit consists of D-galactose (β1⟶4)-linked to 3,6-anhydro-L-galactose (in which an ether bridge connects C-3 and C-6). These units are joined by (α1⟶3) glycosidic links to form a polymer 600 to 700 residues long. A small fraction of the 3,6-anhydrogalactose residues have a sulfate ester at C-2 (as shown here). The open parentheses in the systematic name indicate that the repeating unit extends from both ends.

lysozyme kills bacteria by hydrolyzing the (β1⟶4) glycosidic bond between *N*-acetylglucosamine and *N*-acetylmuramic acid (see Fig. 6–27); the enzyme is found in human tears, where it is presumably a defense against bacterial infections of the eye, and is also produced by certain bacterial viruses to ensure their release from the host bacterial cell, an essential step of the viral infection cycle. Penicillin and related antibiotics kill bacteria by preventing synthesis of the cross-links, leaving the cell wall too weak to resist osmotic lysis (p. 224).

Certain marine red algae, including some of the seaweeds, have cell walls that contain **agar**, a mixture of sulfated heteropolysaccharides made up of D-galactose and an L-galactose derivative ether-linked between C-3 and C-6. Agar is a complex mixture of polysaccharides, all with the same backbone structure but substituted to varying degrees with sulfate and pyruvate. **Agarose** (M_r ~150,000) is the agar component with the fewest charged groups (sulfates, pyruvates) **(Fig. 7–21)**. The remarkable gel-forming property of agarose makes it useful in the biochemistry laboratory. When a suspension of agarose in water is heated and cooled, the agarose forms a double helix: two molecules in parallel orientation twist together with a helix repeat of three residues; water molecules are trapped in the central cavity. These structures in turn associate with each other to form a gel—a three-dimensional matrix that traps large amounts of water. Agarose gels are used as inert supports for the electrophoretic separation of nucleic acids, an essential part of the DNA-sequencing process (p. 302). Agar is also used to form a surface for the growth of bacterial colonies. Another commercial use of agar is for the capsules in which some vitamins and drugs are packaged; the dried agar material dissolves readily in the stomach and is metabolically inert.

Glycosaminoglycans Are Heteropolysaccharides of the Extracellular Matrix

The extracellular space in the tissues of multicellular animals is filled with a gel-like material, the **extracellular matrix (ECM)**, also called ground substance, which holds the cells together and provides a porous pathway for the diffusion of nutrients and oxygen to individual cells. The ECM that surrounds fibroblasts and other connective tissue cells is composed of an interlocking meshwork of heteropolysaccharides and fibrous proteins such as fibrillar collagens, elastins, and fibronectins. Basement membrane is a specialized ECM that underlies epithelial cells; it consists of specialized collagens, laminins, and heteropolysaccharides. These heteropolysaccharides, the **glycosaminoglycans**, are a family of linear polymers composed of repeating disaccharide units **(Fig. 7–22)**. They are unique to animals and bacteria and are not found in plants. One of the two monosaccharides is always either *N*-acetylglucosamine or *N*-acetylgalactosamine; the other is in most cases a uronic acid, usually D-glucuronic or L-iduronic acid. Some glycosaminoglycans contain esterified sulfate groups. The combination of sulfate groups and the carboxylate groups of the uronic acid residues gives glycosaminoglycans a very high density of negative charge. To minimize the repulsive forces among neighboring charged groups, these molecules assume an extended conformation in solution, forming a rodlike helix in which the negatively charged carboxylate groups occur on alternate sides of the helix (as shown for heparin in Fig. 7–22). The extended rod form also provides maximum separation between the negatively charged sulfate groups. The specific patterns of sulfated and nonsulfated sugar residues in glycosaminoglycans provide for specific recognition by a variety of protein ligands that bind electrostatically to these molecules. The sulfated glycosaminoglycans are attached to extracellular proteins to form proteoglycans (Section 7.3).

The glycosaminoglycan **hyaluronan** (hyaluronic acid) contains alternating residues of D-glucuronic acid and *N*-acetylglucosamine (Fig. 7–22). With up to 50,000 repeats of the basic disaccharide unit, hyaluronan has a

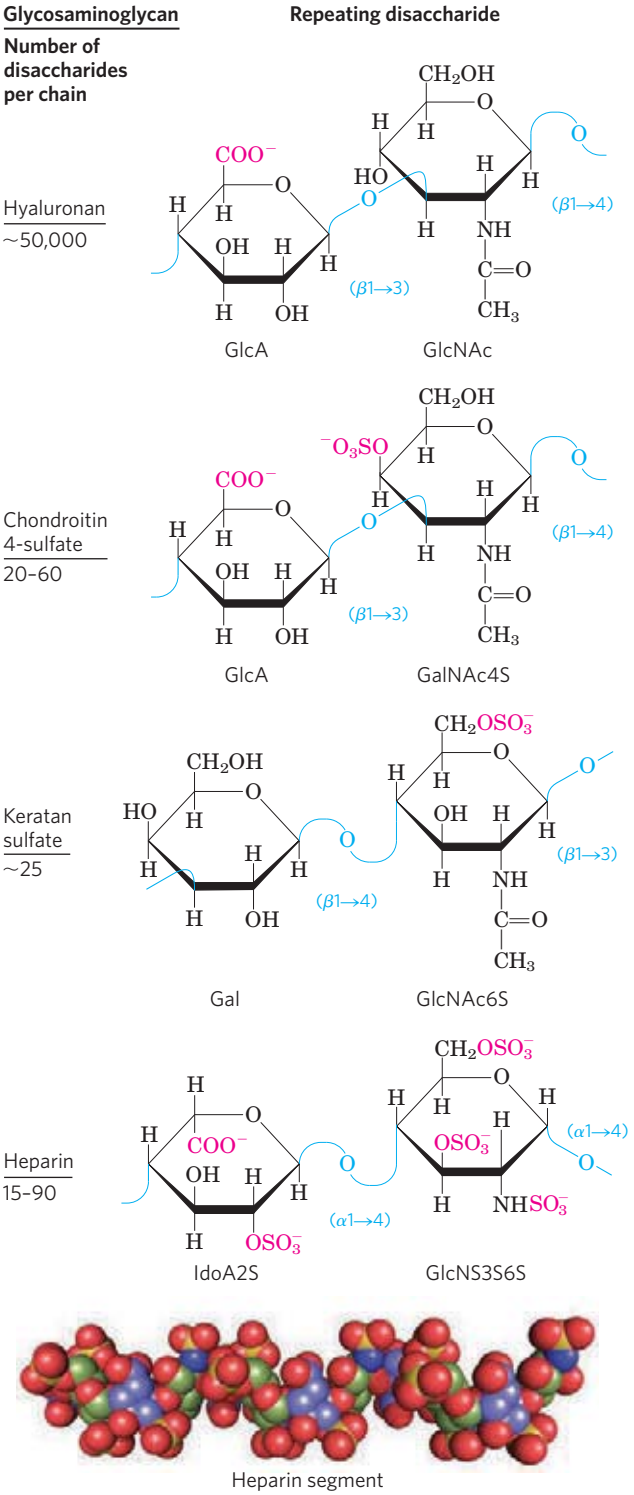
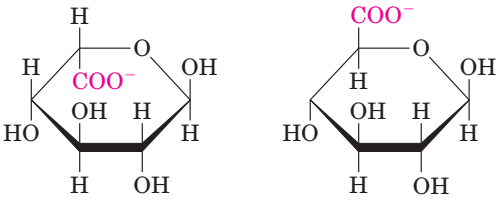


FIGURE 7–22 Repeating units of some common glycosaminoglycans of extracellular matrix. The molecules are copolymers of alternating uronic acid and amino sugar residues (keratan sulfate is the exception), with sulfate esters in any of several positions, except in hyaluronan. The ionized carboxylate and sulfate groups (red in the perspective formulas) give these polymers their characteristic high negative charge. Therapeutic heparin contains primarily iduronic acid (IdoA) and a smaller proportion of glucuronic acid (GlcA, not shown), and is generally highly sulfated and heterogeneous in length. The space-filling model shows a heparin segment as its solution structure, as determined by NMR spectroscopy (PDB ID 1HPN). The carbons in the iduronic acid sulfate are colored blue; those in glucosamine sulfate are green. Oxygen and sulfur atoms are shown in their standard colors of red and yellow, respectively. The hydrogen atoms are not shown (for clarity). Heparan sulfate (not shown) is similar to heparin but has a higher proportion of GlcA and fewer sulfate groups, arranged in a less regular pattern.

molecular weight of several million; it forms clear, highly viscous solutions that serve as lubricants in the synovial fluid of joints and give the vitreous humor of the vertebrate eye its jellylike consistency (the Greek *hyalos* means "glass"; hyaluronan can have a glassy or translucent appearance). Hyaluronan is also a component of the extracellular matrix of cartilage and tendons, to which it contributes tensile strength and elasticity as a result of its strong noncovalent interactions with other components of the matrix. Hyaluronidase, an enzyme secreted by some pathogenic bacteria, can hydrolyze the glycosidic linkages of hyaluronan, rendering tissues more susceptible to bacterial invasion. In many animal species, a similar enzyme in sperm hydrolyzes an outer glycosaminoglycan coat around the ovum, allowing sperm penetration.

Other glycosaminoglycans differ from hyaluronan in three respects: they are generally much shorter polymers, they are covalently linked to specific proteins (proteoglycans), and one or both monomeric units differ from those of hyaluronan. **Chondroitin sulfate** (Greek *chondros*, "cartilage") contributes to the tensile strength of cartilage, tendons, ligaments, and the walls of the aorta. Dermatan sulfate (Greek *derma*, "skin") contributes to the pliability of skin and is also present in blood vessels and heart valves. In this polymer, many of the glucuronate residues present in chondroitin sulfate are replaced by their 5-epimer, L-iduronate (IdoA).



Keratan sulfates (Greek *keras*, "horn") have no uronic acid and their sulfate content is variable. They are present in cornea, cartilage, bone, and a variety of horny structures formed of dead cells: horn, hair, hoofs, nails, and claws. **Heparan sulfate** (Greek *hēpar*, "liver"; it was originally isolated from dog liver) is produced by all animal cells and contains variable arrangements of sulfated and nonsulfated sugars. The sulfated segments of the chain allow it to interact with a large number of proteins, including growth factors and ECM components, as well as various enzymes and factors present in plasma. Heparin is a fractionated form of heparan sulfate derived mostly from mast cells (a type of leukocyte). Heparin is a therapeutic agent used to inhibit coagulation through its capacity to bind the protease inhibitor antithrombin.

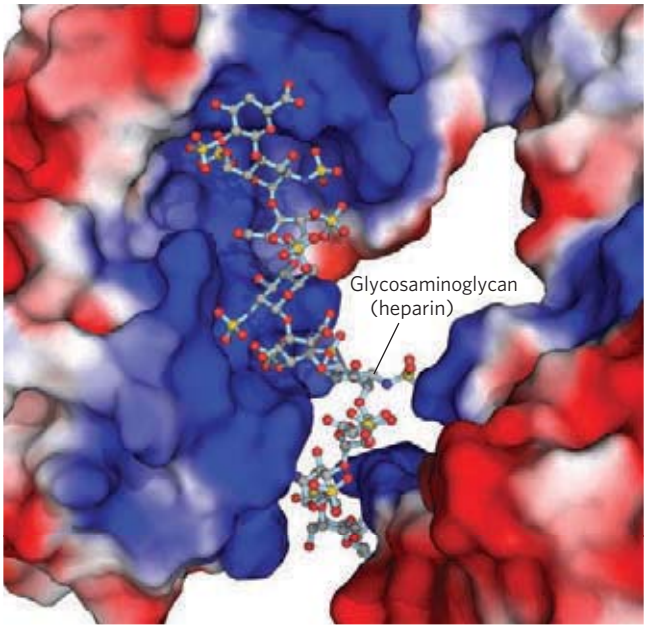


FIGURE 7–23 Interaction between a glycosaminoglycan and its binding protein. Fibroblast growth factor 1 (FGF1), its cell surface receptor (FGFR), and a short segment of a glycosaminoglycan (heparin) were co-crystallized to yield the structure shown here (PDB ID 1E0O). The proteins are represented as surface contour images, with color to represent surface electrostatic potential: red, predominantly negative charge; blue, predominantly positive charge. Heparin is shown in a ball-and-stick representation, with the negative charges ($—SO_3^-$ and $—COO^-$) attracted to the positive (blue) surface of the FGF1 protein. Heparin was used in this experiment, but the glycosaminoglycan that binds FGF1 in vivo is heparan sulfate on the cell surface.

Heparin binding causes antithrombin to bind to and inhibit thrombin, a protease essential to blood clotting. The interaction is strongly electrostatic; heparin has the highest negative charge density of any known biological macromolecule **(Fig. 7–23)**. Purified heparin is routinely added to blood samples obtained for clinical analysis, and to blood donated for transfusion, to prevent clotting.

Table 7–2 summarizes the composition, properties, roles, and occurrence of the polysaccharides described in Section 7.2.

TABLE 7–2 Structures and Roles of Some Polysaccharides

Polymer	Type*	Repeating unit†	Size (number of monosaccharide units)	Roles/significance
Starch				Energy storage: in plants
Amylose	Homo-	(α1→4)Glc, linear	50–5,000	
Amylopectin	Homo-	(α1→4)Glc, with (α1→6)Glc branches every 24–30 residues	Up to 10^6	
Glycogen	Homo-	(α1→4)Glc, with (α1→6)Glc branches every 8–12 residues	Up to 50,000	Energy storage: in bacteria and animal cells
Cellulose	Homo-	(β1→4)Glc	Up to 15,000	Structural: in plants, gives rigidity and strength to cell walls
Chitin	Homo-	(β1→4)GlcNAc	Very large	Structural: in insects, spiders, crustaceans, gives rigidity and strength to exoskeletons
Dextran	Homo-	(α1→6)Glc, with (α1→3) branches	Wide range	Structural: in bacteria, extracellular adhesive
Peptidoglycan	Hetero-; peptides attached	4)Mur2Ac(β1→4)GlcNAc(β1	Very large	Structural: in bacteria, gives rigidity and strength to cell envelope
Agarose	Hetero-	3)D-Gal(β1→4)3,6-anhydro-L-Gal(α1	1,000	Structural: in algae, cell wall material
Hyaluronan (a glycosaminoglycan)	Hetero-; acidic	4)GlcA(β1→3)GlcNAc(β1	Up to 100,000	Structural: in vertebrates, extracellular matrix of skin and connective tissue; viscosity and lubrication in joints

*Each polymer is classified as a homopolysaccharide (homo-) or heteropolysaccharide (hetero-).

†The abbreviated names for the peptidoglycan, agarose, and hyaluronan repeating units indicate that the polymer contains repeats of this disaccharide unit. For example, in peptidoglycan, the GlcNAc of one disaccharide unit is (β1→4)-linked to the first residue of the next disaccharide unit.

SUMMARY 7.2 Polysaccharides

- Polysaccharides (glycans) serve as stored fuel and as structural components of cell walls and extracellular matrix.
- The homopolysaccharides starch and glycogen are stored fuels in plant, animal, and bacterial cells. They consist of D-glucose with (α1→4) linkages, and both contain some branches.
- The homopolysaccharides cellulose, chitin, and dextran serve structural roles. Cellulose, composed of (β1→4)-linked D-glucose residues, lends strength and rigidity to plant cell walls. Chitin, a polymer of (β1→4)-linked *N*-acetylglucosamine, strengthens the exoskeletons of arthropods. Dextran forms an adhesive coat around certain bacteria.
- Homopolysaccharides fold in three dimensions. The chair form of the pyranose ring is essentially rigid, so the conformation of the polymers is determined by rotation about the bonds from the rings to the oxygen atom in the glycosidic linkage. Starch and glycogen form helical structures with intrachain hydrogen bonding; cellulose and chitin form long, straight strands that interact with neighboring strands.
- Bacterial and algal cell walls are strengthened by heteropolysaccharides—peptidoglycan in bacteria, agar in red algae. The repeating disaccharide in peptidoglycan is GlcNAc(β1→4)Mur2Ac; in agar, it is D-Gal(β1→4)3,6-anhydro-L-Gal.
- Glycosaminoglycans are extracellular heteropolysaccharides in which one of the two monosaccharide units is a uronic acid (keratin sulfate is an exception) and the other an *N*-acetylated amino sugar. Sulfate esters on some of the hydroxyl groups and on the amino group of some glucosamine residues in heparin and in heparan sulfate give these polymers a high density of negative charge, forcing them to assume extended conformations. These polymers (hyaluronan, chondroitin sulfate, dermatan sulfate, and keratan sulfate) provide viscosity, adhesiveness, and tensile strength to the extracellular matrix.

7.3 Glycoconjugates: Proteoglycans, Glycoproteins, and Glycosphingolipids

In addition to their important roles as stored fuels (starch, glycogen, dextran) and as structural materials (cellulose, chitin, peptidoglycans), polysaccharides and oligosaccharides are information carriers. Some provide communication between cells and their extracellular surroundings; others label proteins for transport to and localization in specific organelles, or for destruction when the protein is malformed or superfluous; and others serve as recognition sites for extracellular signal molecules (growth factors, for example) or extracellular parasites (bacteria or viruses). On almost every eukaryotic cell, specific oligosaccharide chains attached to components of the plasma membrane form a carbohydrate layer (the glycocalyx), several nanometers thick, that serves as an information-rich surface that the cell shows to its surroundings. These oligosaccharides are central players in cell-cell recognition and adhesion, cell migration during development, blood clotting, the immune response, wound healing, and other cellular processes. In most of these cases, the informational carbohydrate is covalently joined to a protein or a lipid to form a **glycoconjugate**, which is the biologically active molecule **(Fig. 7–24)**.

Proteoglycans are macromolecules of the cell surface or extracellular matrix in which one or more sulfated glycosaminoglycan chains are joined covalently to a membrane protein or a secreted protein. The glycosaminoglycan chain can bind to extracellular proteins through electrostatic interactions between the protein and the negatively charged sugar moieties on the proteoglycan. Proteoglycans are major components of all extracellular matrices.

Glycoproteins have one or several oligosaccharides of varying complexity joined covalently to a protein. They are usually found on the outer face of the plasma membrane (as part of the glycocalyx), in the extracellular matrix, and in the blood. Inside cells they are found in specific organelles such as Golgi complexes,

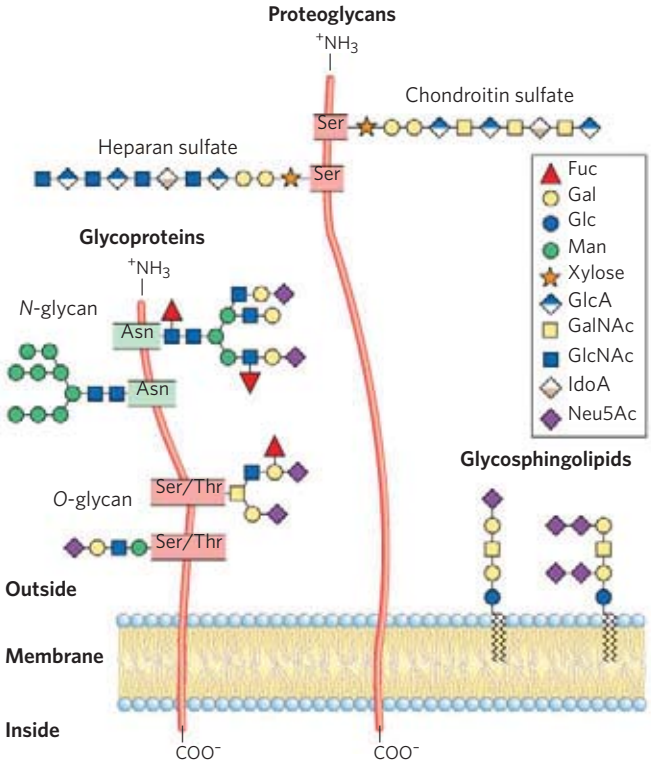


FIGURE 7–24 Glycoconjugates. The structures of some typical proteoglycans, glycoproteins, and glycosphingolipids described in the text.