

Carbohydrates

Learning Objectives

1. Define and use correctly the terms *aldose*, *ketose*, *pentose*, *hexose*, *diastereomer*, *enantiomer*, *pyranose*, *furanose*, *anomer*, *glycoside*, *starch*, *amylose*, *cellulose*, *glycosaminoglycan*, *proteoglycan*, *glycolipid*, *glycoprotein*, and *lectin*.
2. Recognize the structures of glucose, fructose, glucuronic acid, glucosamine, and *N*-acetylglucosamine.
3. Explain the *D*, *L* nomenclature system for simple carbohydrates; note that it does not predict the direction of rotation of plane polarized light.
4. Explain what is meant by the term *reducing sugar*. Distinguish between the reducing and nonreducing ends of a polysaccharide such as glycogen.
5. List several common biologically relevant modifications of simple sugars.
6. List three common disaccharides and their natural sources. Describe how they differ in structure.
7. Describe the structure of common simple polysaccharides such as starch, cellulose, and glycogen.
8. Describe the structure of common glycosaminoglycans such as hyaluronic acid, heparan sulfate, and chondroitin sulfate. Describe how they may be linked to proteins to form proteoglycans. Relate chemical features of these polysaccharides to the biological function of these polymers.
9. Describe the general biosynthetic pathway leading to mature proteoglycans, involving the endoplasmic reticulum and the Golgi apparatus.
10. List several common glycoproteins and give their function.
11. Explain how sugars are linked to proteins via asparagine and serine or threonine.
12. Describe differences in the glycolipids, especially the carbohydrate content, contributing to the different blood group antigens. Explain the biochemical basis for these differences.
13. Describe how elevated blood sugar can lead to hemoglobin glycosylation, and how quantitation of HbA_{1c} is used in following the effectiveness of treatment of diabetes.
14. Relate the release of heparan sulfate after tissue injury to its effects on blood clotting. Identify the differences between heparan sulfate and heparin.

Simple Sugars: Monosaccharides

Introduction

Carbohydrates (sugars and polymers made of sugars) are ubiquitous in biology and play many important roles. Here are just a few of the many ways carbohydrates are basic to biology:

- The simple sugar glucose is a principal source of metabolic energy for the cell. Cellulose, a polymer of glucose, strengthens the cell walls of plants.
- The cell uses carbohydrate chains to modify proteins and direct those proteins to their proper sites for function in the cell.
- In animals, the complex polymers known as glycosaminoglycans (chains of unmodified and modified sugars) lubricate and cushion joints; help form skin, hair, and feathers; help regulate blood clotting and assist the immune system; and help control cell-to-cell adhesion and interaction.

Basic Nomenclature

The formula for a carbohydrate, $(\text{CH}_2\text{O})_n$, may be simple, but there is a great deal of complexity in the structures of such compounds. Inevitably, to properly describe this structural diversity, there is also a great deal of detailed nomenclature.

Conventionally, a monosaccharide is a carbohydrate that does not hydrolyze, while a disaccharide can be hydrolyzed to give two monosaccharides, a trisaccharide can be hydrolyzed to give three monosaccharides, and so on. A polysaccharide contains many monosaccharide units. Also, the smallest molecules regarded as carbohydrates have three carbon atoms; the exemplars here are glyceraldehyde and dihydroxyacetone (**Figure 2-1**). Compounds with only one or two carbons do not share extensive properties with those we regard as “sugars,” so they are not regarded as saccharides.

The suffix “-ose” on a compound’s name indicates a sugar. An *aldose* is a sugar derived from an aldehyde, and a *ketose* is derived from a ketone. A *triose* is a carbohydrate with three carbon atoms; a *tetrose* has four carbons; a *pentose* has five carbons; and a *hexose* has six carbons. These terms may be combined for a more complete specification of the compound. For example, an aldopentose is a sugar with five carbons and an aldehyde group.

Stereochemistry

A very important feature of sugars is the large number of isomers possible within the overall molecular formula, even for simple pentoses and hexoses. A quick review of some concepts and terminology from organic chemistry will help to keep matters clear.

- *Stereoisomers* are molecules that have the same molecular formula but differ in the way that the constituent atoms are oriented in space.

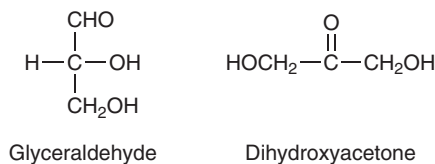


Figure 2-1 The smallest saccharides, glyceraldehyde and dihydroxyacetone.

- *Enantiomers* are molecules that contain a center of asymmetry, usually a carbon atom known as a chiral (“handed”) carbon. Enantiomers are mirror images of each other. They have the same physical properties, except they rotate polarized light in different directions. By convention, a + sign indicates rotation of the polarization to the right, while a – sign indicates rotation to the left. Because their physical properties are the same, it is often very difficult to physically separate enantiomers from each other.
- *Diastereomers* are stereoisomers that are not mirror images. Their physical properties differ, and this permits their physical separation. Note that a pair of stereoisomers that are not enantiomers must then be diastereomers.

Glyceraldehyde contains a chiral carbon, and there are two possible enantiomers of this compound. **Figure 2-2** illustrates these two forms using a Fischer projection. By convention, the C1 carbon is the aldehydic carbon, and it is drawn at the top of a vertical line of carbon atoms, C1 through C3. The chiral carbon is C2; the OH group is drawn to the right of the chiral carbon for what is denoted as the D isomer; the L isomer has this group drawn to the left side of this carbon.

In terms of visualizing the direction of bonds around C2, the vertically aligned bonds are imagined to be going into the plane of the drawing and away from the viewer, while the horizontally aligned bonds are directed upward, out of the plane of the drawing and toward the viewer. **Figure 2-3** presents the same structures using wedges in place of lines, to give a better sense of perspective on the direction of the bonds.

With respect to rotation of plane polarized light, the D isomer here rotates it to the right (the D originally indicated *dextro*, or right-handed, rotation and the L indicated *levo*, or left-handed, rotation). The D, L convention for stereoisomers can also be applied to other simple chiral compounds, such as amino acids. However, while many D monosaccharides do, indeed, rotate plane-polarized light to the right, this behavior is not universal. Instead, to indicate this particular physical property, as distinct from the configuration about the chiral carbon, it is preferable to use a prefix of + or –, with the + sign indicating right-handed rotation and the – sign indicating left-handed rotation.

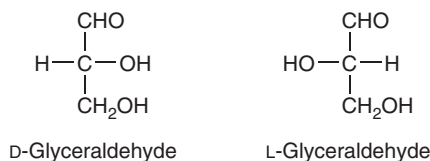


Figure 2-2 D- and L-glyceraldehyde as Fischer projections.

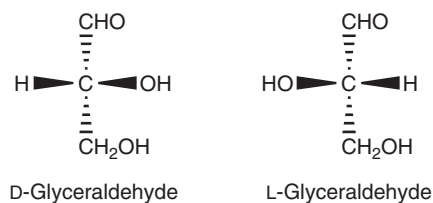


Figure 2-3 D- and L-glyceraldehyde drawn using “wedge” bonds.

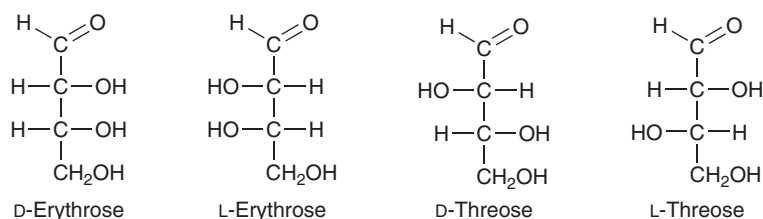


Figure 2-4 Pairs of enantiomers for the aldotetroses: D- and L-erythrose, and D- and L-threose.

Tetroses have four carbons, two of which may be chiral. Consider first the possibilities for the aldotetroses. **Figure 2-4** shows erythrose and threose; there are mirror images for each of these compounds, so there are four stereoisomers overall—that is, two pairs of enantiomers. Note that either enantiomer of erythrose is a diastereomer of either enantiomer of threose. Either threose or erythrose can be tautomerized to a ketose, erythulose; D-erythrose gives D-erythulose, which in turn gives D-threose (**Figure 2-5**). A similar pattern holds for the interconversion of the L isomers.

With aldopentoses there are three chiral carbons, giving four different pairs of enantiomers: D and L forms for ribose, arabinose, xylose, and lyxose (**Figure 2-6**). There are also two ketopentoses, ribulose and xylulose, each with a D and an L form. The D isomers are the

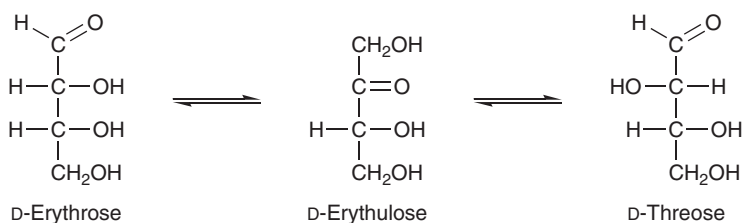


Figure 2-5 Tautomerization of D-erythrose to D-erythulose, and then to D-threose.

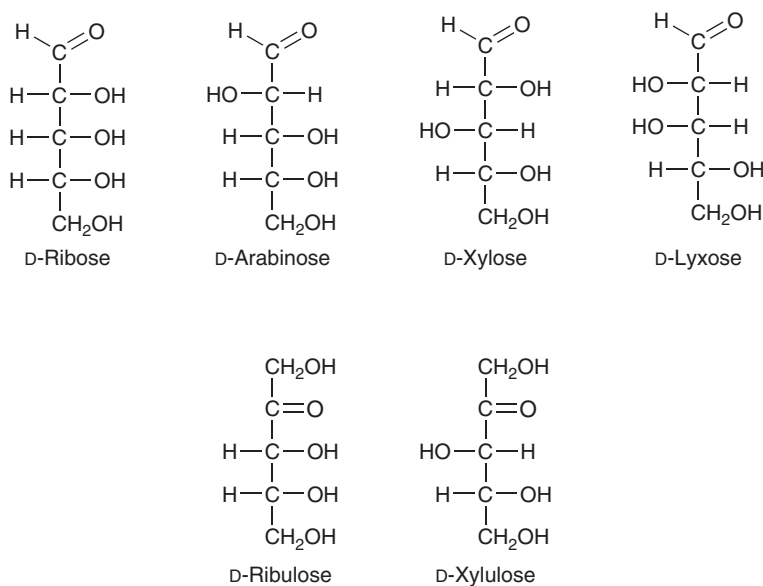


Figure 2-6 Pentoses. Only the D isomers are shown.

most important for biochemistry. Additionally, the important pentose derived from ribose, deoxyribose, is essential to the structure of DNA.

There are 4 chiral carbons for aldohexoses and 16 diastereomers. D-Glucose, D-mannose, and D-galactose are the most important for human biochemistry. There are also four ketohexoses, with D-fructose being the most important species. D-(+)-Glucose is sometimes referred to as dextrose, while levulose is D-(−)-fructose.

Cyclization

The aldopentoses and the hexoses show a strong tendency to cyclize in solution, forming rings with five or six members. These rings contain an oxygen atom as one of the members. Thus sugars with a five-membered ring resemble the compound furan, while those with a six-membered ring resemble pyran (**Figure 2-7**). This leads to the practice of designating the sugars in such conformations as *furanoses* and *pyranoses*, respectively. As a pyranose, glucose forms a hemiacetal (the product of a 1:1 reaction of an aldehyde with an alcohol; see **Figure 2-8**), which is relatively stable thanks to its cyclic structure. This reaction creates a new chiral center at C1 and leads to two new optically active forms, called *anomers*; C1 is referred to as the anomeric carbon in this case. The anomers are designated as α -D-glucopyranose and β -D-glucopyranose (**Figure 2-9**); these names are often abbreviated simply to α -D-glucose and β -D-glucose. Figure 2-9 shows these anomers in the stable “chair” conformation. Here the —OH groups at C2, C3, C4, and C5 are equatorial. However, notice that the —OH group at C1 may be either equatorial (the β anomer) or axial (the α anomer).

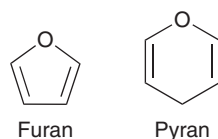


Figure 2-7 Furan and pyran.

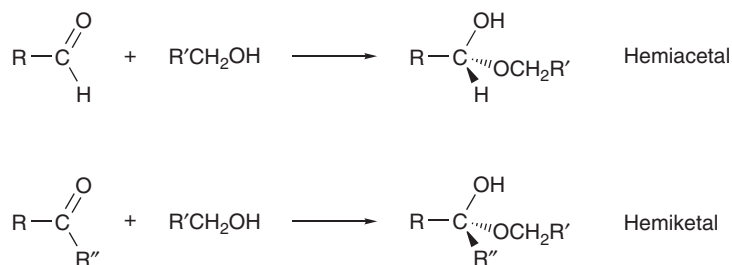


Figure 2-8 Hemiacetal and hemiketal formation.

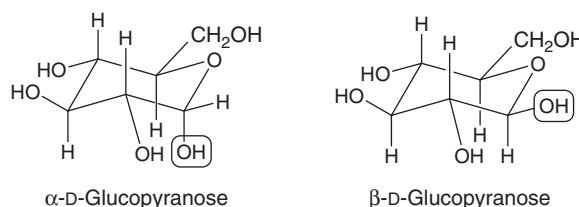


Figure 2-9 Anomers of D-glucose. Some bond lengths have been extended to emphasized structural details and the —OH group attached to the anomeric carbon has been marked.

Fructose can cyclize to give either a furanose ring or a pyranose, through formation of a hemiketal (the product of a nucleophilic alcohol reacting with the carbonyl of a ketone). For either the pyranose or the furanose, two anomers are possible, α - and β -D-fructofuranose and α - and β -D-fructopyranose (**Figure 2-10**). In solution, the pyranose form predominates (approximately 60%) over the furanose form (approximately 40%), with a tiny amount of the open-chain form (much less than 1%). Ribose and deoxyribose (**Figure 2-11**) can also cyclize to form a furanose, with two anomers.

Reactivity and Common Derivatives

The hemiacetals and hemiketals formed in cyclizing pentoses and hexoses are still in equilibrium with the open-chain form of these sugars, although the equilibrium generally favors the cyclized form by a considerable margin. As a consequence, in solutions of these sugars, there will be small amounts of “free” aldehyde or ketone available for reaction. Aldehydes, such as aldoses, can be oxidized by mild agents. This behavior is the basis for the classic organic chemistry lab tests for an aldehyde, such as the silver mirror produced by Tollens’ reagent, or the color changes produced by Benedict’s solution or Fehling’s solution. Because solutions of aldoses and ketoses will have small amounts of the “free” aldehyde or ketone, they can also react with these reagents. Sugars that give positive test results with these reagents are termed reducing sugars. The products of the oxidation of aldoses are called aldonic acids.

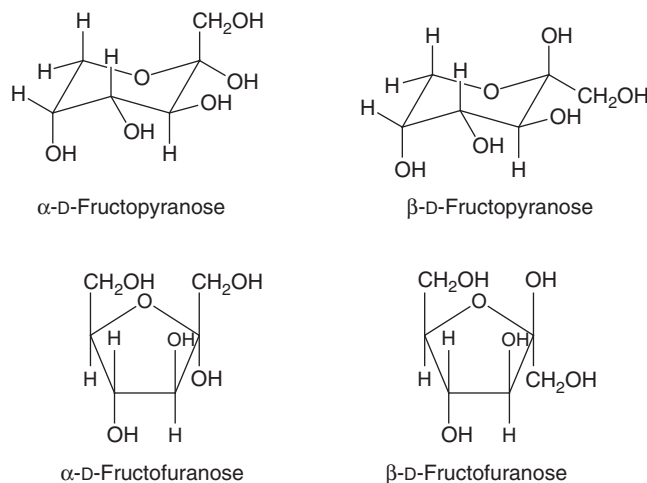


Figure 2-10 Anomers of D-fructose as a pyranose and furanose.

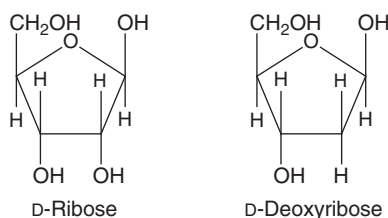


Figure 2-11 D-Ribose and D-deoxyribose.

Conversely, the carbonyl group on aldoses and ketoses may be reduced to give the corresponding alcohol. **Figure 2-12** shows two important examples of this class of compounds. D-Sorbitol is also called D-glucitol. In the disease of diabetes mellitus, sorbitol accumulates in the lens of the eye and is associated with the formation of cataracts. The mechanism of cataract formation may possibly involve redox reactions in the formation of sorbitol from glucose and associated oxidative cross-linking of proteins through disulfide bridges.

Enzymes can oxidize monosaccharides to give a number of different products. Two important carboxylic acids that are derived in this way are *glucuronic acid* and *iduronic acid* (**Figure 2-13**). Note that iduronic acid is an epimer of glucuronic acid; an epimer is one of a pair of diastereomeric aldoses that differ only at the C2 position in their configuration.

Glycosides are formed from hemiacetals or hemiketals (that is, reducing sugars) by reaction with an alcohol. The product is a full acetal or full ketal (**Figure 2-14**). Glycosides are not reducing sugars. In naming these compounds, “glyc” is the general prefix for the combination

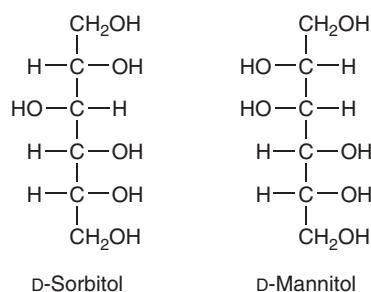


Figure 2-12 Sorbitol and mannitol are two important alditols, derived from the reduction of the corresponding aldoses.

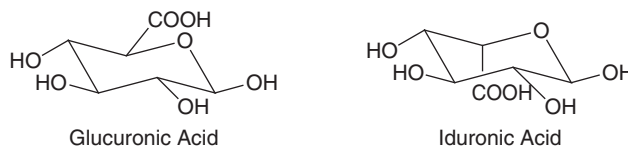


Figure 2-13 Uronic acids: glucuronic and iduronic acid.

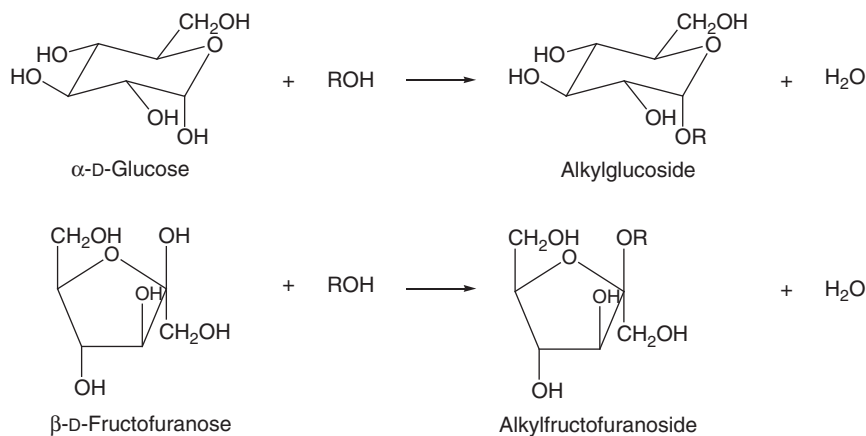


Figure 2-14 Glycoside formation, giving a full acetal or full ketal.

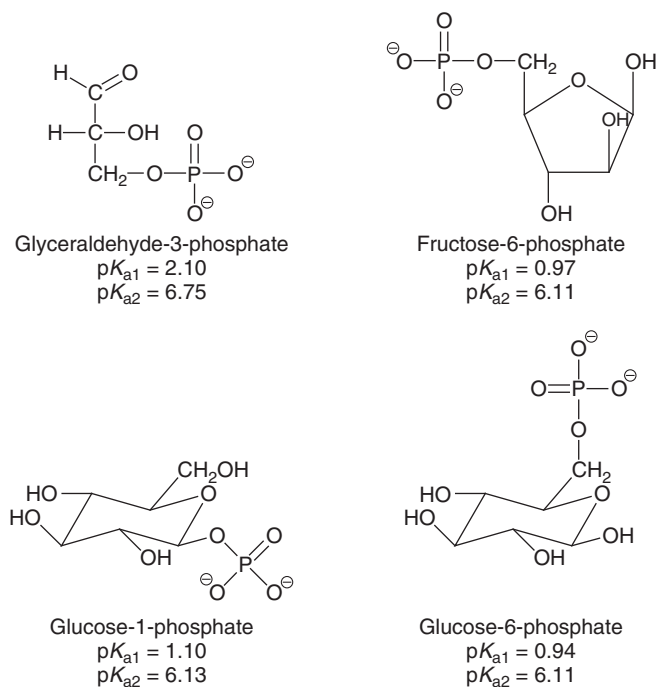


Figure 2-15 Sugar–phosphate esters and their pK_a values.

of an unspecified sugar and alcohol. Thus reaction with methanol would give a methylglycoside, for example. If the sugar is specifically identified, one may have a glucoside (from glucose), a galactoside (from galactose), and so on. Glycosides are fairly stable and are not readily cleaved by hydroxide ion. However, the bond linking the alcohol and sugar can be cleaved by enzymes called glycosidases. Glycosides are widely distributed in plants and animals.

As alcohols, sugars can react with phosphoric acid to form phosphate esters. These compounds are quite important in the metabolism of carbohydrates, forming many important intermediates. Sugar–phosphate esters are acidic, even more so than orthophosphoric acid. **Figure 2-15** collects several important examples of carbohydrate phosphate esters that are important in energy metabolism. Hydrolysis of these compounds is thermodynamically favorable, a point that we will return to later (Chapter 9) as we discuss carbohydrate metabolism.

Natural polysaccharides may contain sugars that have been modified by sulfation or attachment of an amino group. Two common amino sugars are β -D-glucosamine and β -D-galactosamine (**Figure 2-16**). These amino sugars can be further modified by acetylation or other alterations. Also, in these natural polysaccharides, sulfate groups may be attached to sugars in a variety of ways (**Figure 2-17**). When dealing with these more complex compounds, it is convenient to use abbreviations for the monosaccharides and their derivatives. **Table 2-1** collects some of the most commonly used abbreviations.

Oligosaccharides and Polysaccharides

The terms “oligomer” and “polymer” are used to indicate molecules composed of linked units of simple constitution. The simple units are called monomers, and a molecule containing many monomers is called a polymer. An oligomer is essentially a short polymer; it contains two or more monomer units, but not as many as a polymer. The size difference between a polymer and an oligomer is not well defined. It is safe to say that a chain of 100 monomer units would

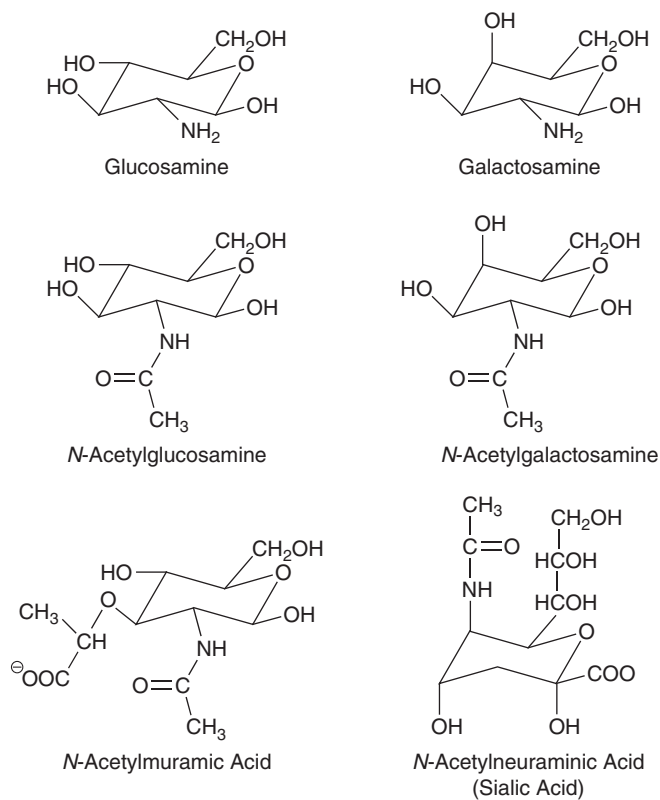


Figure 2-16 Amino sugars and their derivatives.

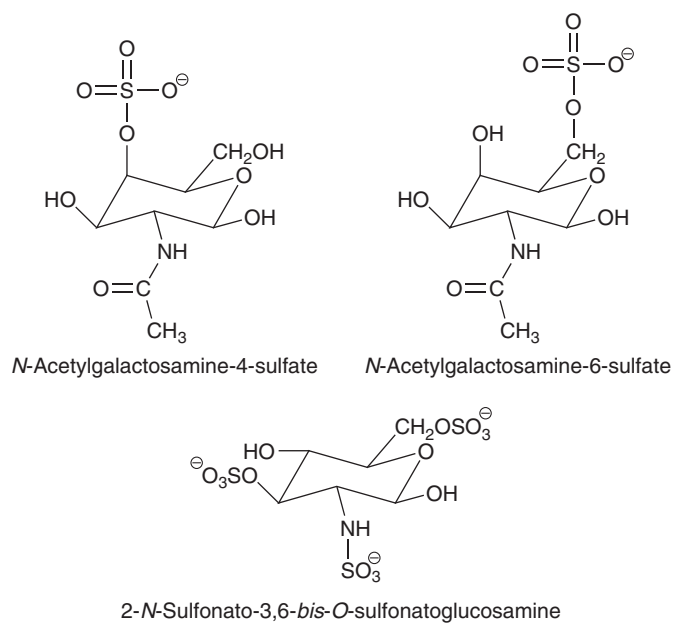


Figure 2-17 Sulfated sugars found in natural polysaccharides.

Monosaccharide	Abbreviation
Glucose	Glc
Fructose	Fru
Galactose	Gal
Mannose	Man
Xylose	Xyl
Fucose	Fuc
β -D-acetylgalactosamine	GalNAc
β -D-acetylglucosamine	GlcNAc
Glucuronic acid	GlcA
Iduronic acid	IdoA
Sialic acid	Sia

Table 2-1 Common Simple and Modified Monosaccharides Found in Glycosaminoglycans

qualify as a polymer, and a chain of only 10 or 20 units would probably be classified as an oligomer, but a chain of 50 units might in some cases be referred to as a polymer, and in other instances as an oligomer.

The prefixes “poly-” and “oligo-” can be combined with the name of the type of monomeric unit to specify the type of polymer or oligomer. For example, a long chain whose repeating monomeric units are all sugars would be regarded as a polysaccharide; a short chain of two or three nucleic acid units (nucleotides) might called an oligonucleotide, and so on.

Oligosaccharides

There are three common disaccharides: *sucrose*, *lactose*, and *maltose*. Sucrose, or common table sugar, is commercially prepared from sugar cane or sugar beets. It is widely found in plants. Sucrose itself is not a reducing sugar, so there are no “free” hemiacetal or hemiketal moieties in the disaccharide. Upon hydrolysis, sucrose gives D-(+)-glucose and D-(-)-fructose. These monomers are linked through the anomeric carbons, C1 of glucose and C2 of fructose (we shall abbreviate this as a 1 \rightarrow 2 linkage), as shown in **Figure 2-18**, to form α -D-glucopyranosyl- β -D-fructofuranoside, which is both a glucoside and a fructoside.

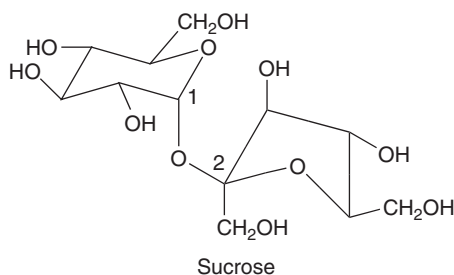


Figure 2-18 Sucrose: α -D-glucopyranosyl- β -D-fructofuranoside.

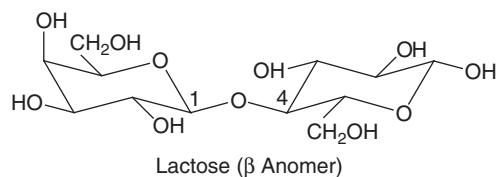


Figure 2-19 β -Lactose: 4- β -D-galactopyranosyl-D-glucopyranose.

Lactose is present in milk at approximately 5% concentration and is commercially prepared as a dairy by-product. A disaccharide of galactose and glucose, it is itself a reducing sugar, with the glucose attached to one of the oxygens of the galactose unit. This makes lactose technically a galactoside, not a glucoside; it is the glucose unit with its “free” hemiacetal that gives the disaccharide its character as a reducing sugar. Formally, lactose is 4- β -D-galactopyranosyl-D-glucose (**Figure 2-19**).

The breakdown of starch yields maltose. This disaccharide is a reducing sugar, which can be hydrolyzed to two molecules of glucose. The linkage is through an acetal, and there are two possible configurations of the anomeric carbon involved: α -maltose (4- α -glucopyranosyl- α -D-glucopyranose) or β -maltose (4- α -glucopyranosyl- β -D-glucopyranose). See **Figure 2-20**.

In later sections and chapters, we will usually drop the detailed specification of configuration and simply assume the configurations listed here.

Glucose Polymers

Starch is a plant storage form for glucose; it is a polymer of glucose. Acid hydrolysis converts starch to corn syrup. There are two main fractions: water-soluble *amylose* (about 20%) and water-insoluble *amylopectin* (about 80%). Amylose is a linear polymer of around 200 units of glucose with α (1 \rightarrow 4) linkages (**Figure 2-21**). Amylopectin is also a glucose polymer, but of

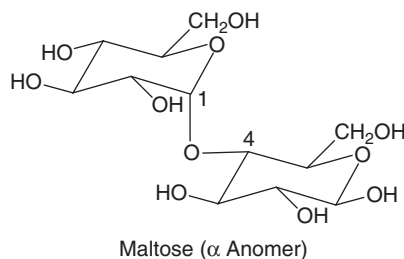


Figure 2-20 α -Maltose: 4- α -glucopyranosyl- α -D-glucopyranose.

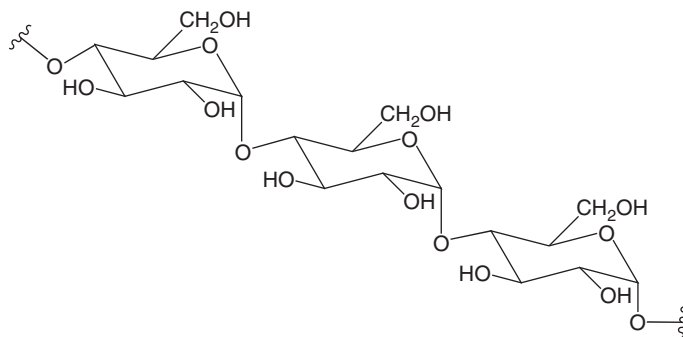


Figure 2-21 Amylose. Hydrogen atoms have been omitted for clarity.

higher molecular weight (at least 1000 glucose units per molecule), with occasional branches in the chain; most of the glucose residues are joined by 1 \rightarrow 4 linkages, but the branches are formed using α (1 \rightarrow 6) linkages.

Glycogen is an animal polymer of glucose, similar to amylopectin, but with more branching. Its synthesis and breakdown are the subject of discussion in Chapter 9. Cellulose is the main polymer of the cell walls of plants; it forms woody structures. It is insoluble in water, unlike starch and glycogen. In cellulose, there are a few thousand glucose units per molecule. Although cellulose is a glucose polymer, it is not digestible by humans or by most animals. The glucose units are joined by β (1 \rightarrow 4) linkages, unlike glycogen or starch (**Figure 2-22**), and animals generally do not make the enzyme needed to hydrolyze these bonds. Ruminant animals such as cattle and insects such as termites carry in their gut certain strains of bacteria that make the enzymes necessary to break down cellulose, enabling their animal hosts to digest cellulose.

As a part of the cell wall, the polymer chains of cellulose lie side-by-side, forming bundles; these bundles are twisted together to create strands and finally visible fibers. Wood contains lignin (cross-linked C9 aromatic units, related in structure and biosynthesis to the aromatic amino acids; **Figure 2-23**) embedded in cellulose fibers; the lignin gives the wood very high strength.

Glycosaminoglycans

Glycosaminoglycans (GAGs) are large linear polymers of repeating disaccharide units, commonly containing one or another amino sugar as one of the monomers in the disaccharide unit. Several types of glycosaminoglycans exist, including chondroitin sulfate, hyaluronan, heparan sulfate, and keratan sulfate; they differ in the type of disaccharide unit used in their synthesis. Glycosaminoglycans are found outside cells, on the cell surface, or as part of the extracellular matrix. They perform various roles, serving as mechanical support and cushioning joints, as cellular signals involved in cell proliferation and cell migration, and as inhibitors of certain key enzymes.

Glycosaminoglycans are often found attached to a protein core, forming what is called a *proteoglycan*. In the past, proteoglycans were called mucopolysaccharides.

Proteoglycans are more carbohydrate than protein, so their properties are mainly determined by the carbohydrate portion of the molecule. These carbohydrate moieties may contain

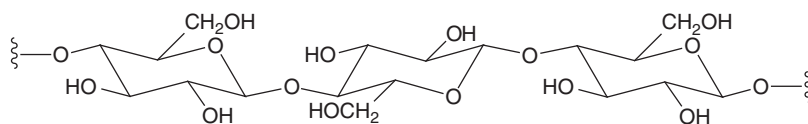


Figure 2-22 Cellulose. Hydrogen atoms have been omitted for clarity.

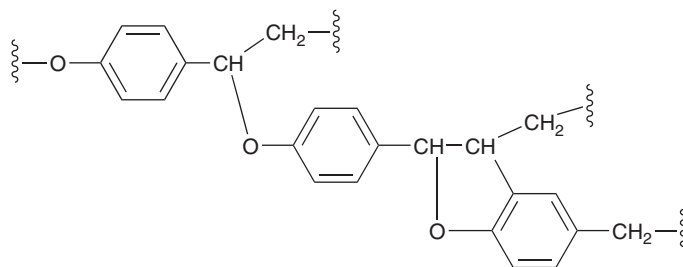


Figure 2-23 Lignin (partial structure).

carboxylic acids (uronic acids) or sulfated sugars, so generally the GAG chains carry a substantial negative charge. Charge repulsions will tend to drive the polysaccharide chains into an extended conformation. Additionally, the charges will help to attract water molecules, so the chains are highly hydrated. The typical proteoglycan is formed from multiple GAG chains on a single polypeptide chain. The overall structure might be described as “feathery”; see **Figure 2-24**.

Heparan sulfate (HS) is a sulfated polysaccharide, found as a component of cell-surface proteoglycans in mast cells, and on the surface of endothelial cells lining blood vessels. It is composed of repeating units of *N*-acetylglucosamine and uronic acids (either glucuronic or iduronic acids). Sulfation (sulfate ester formation) can be found at several positions on these residues; also, the acetyl group on *N*-acetylglucosamine may be replaced by a sulfate group (**Figure 2-25**). After an injury to tissue, oligosaccharides derived from this GAG are released and subsequently help to mediate the inflammatory response. Some of the released oligomers

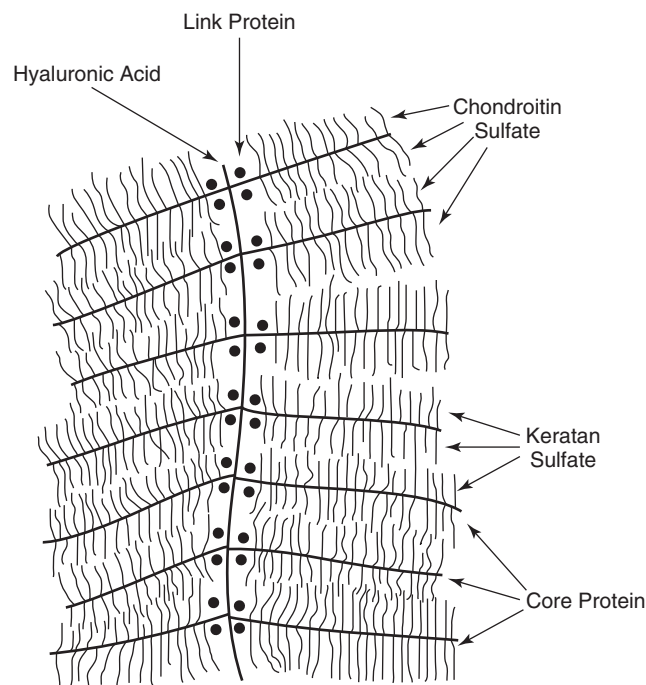


Figure 2-24 Proteoglycan structure. Core protein strands are heavily modified with keratan sulfate and chondroitin sulfate. The core protein strands are held in a complex with a strand of hyaluronic acid by link proteins.

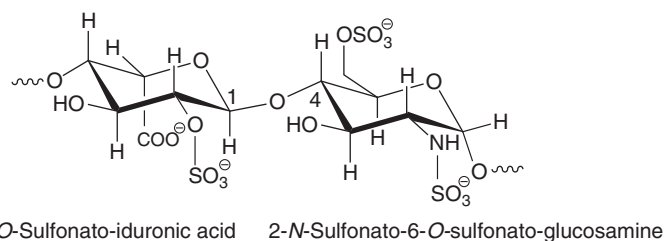


Figure 2-25 Heparan sulfate.

promote activity by growth factors, chemokines, and cytokines; others are associated with recruitment of leukocytes to the injury site. A very important activity is the action of a particular pentasaccharide sequence as an anticoagulant. The fraction of heparan sulfate containing oligomers with this pentasaccharide sequence is designated as *heparin*. The pentasaccharide binds to and activates the enzyme antithrombin III. This enzyme is responsible for inhibiting thrombin, a protease involved in forming blood clots; this inhibition of thrombin, in turn, blocks blood clotting. Note that heparin is much smaller than heparan sulfate and that it is not linked to a protein core. Heparin is also more highly sulfated than the average random pentasaccharide sequence in heparan sulfate. A synthetic version of the key pentasaccharide is used clinically as an anticoagulant (see the discussion later in this chapter in connection with Figure 2-35).

Chondroitin sulfate (CS) is a relatively short polymer, consisting of alternating residues of glucuronic acid and galactose *N*-acetyl 4-sulfonate (**Figure 2-26**). *Dermatan sulfate* (DS) is a closely related GAG, which is composed of glucuronic acid and *N*-acetylgalactosamine. Chondroitin sulfate is found in the extracellular matrix. Its roles are mainly to lend mechanical support and flexibility to tissue; notably, it helps to form skin and cartilage. On membranes, CS and DS are involved in interactions with receptors for growth factors, where they may serve as cofactors for various growth factors.

Keratan sulfate (KS) is found in proteoglycans in three forms, two of which are branched. KS is found primarily in the cornea of the eye and in joint cartilage, where it serves mainly in a mechanical/structural role. It is formed from alternating units of galactose and sulfated *N*-acetylglucosamine (**Figure 2-27**).

Hyaluronic acid (HA; also called hyaluronate) is a long, linear polymer of alternating *N*-acetylglucosamine and glucuronic acid residues (**Figure 2-28**). It serves as a lubricant in joints in the form of synovial fluid; HA is also a principal constituent of the vitreous humor in the eye,

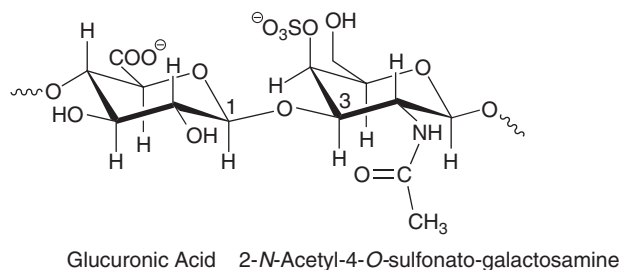


Figure 2-26 Chondroitin sulfate.

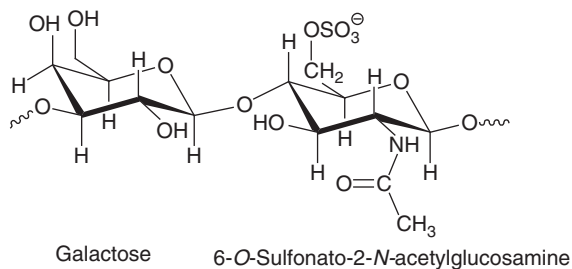


Figure 2-27 Keratan sulfate.

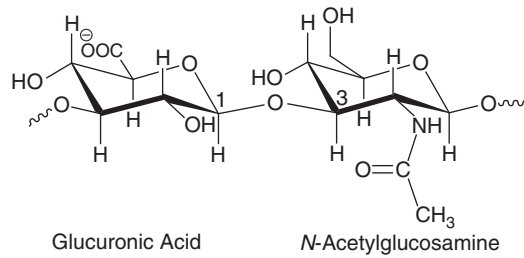


Figure 2-28 Hyaluronic acid.

and it helps to form cartilage. Unlike other GAGs, this polymer is neither sulfated nor attached to a protein core. It is secreted directly to the extracellular matrix. The degradation of HA that occurs after tissue injury releases smaller chains, which can participate in cell proliferation, migration, and differentiation; these degradation products also help to recruit leukocytes to the site of injury. Intra-articular injections of purified HA have been used therapeutically in cases of osteoarthritis.

Other Natural Polysaccharides of Interest

Agar is a linear polymer of sulfated and unsulfated galactose residues, prepared from marine algae; agarose (**Figure 2-29**) is derived from agar as the mostly unsulfated fraction. It is an alternating copolymer of galactose and 3,6-anhydro-galactose. This polymer is not a proteoglycan, but rather is purely carbohydrate. When dissolved in hot water and then cooled, it forms gels with suitable strength for lab use in electrophoresis; it is also used as a food additive to thicken liquid suspensions.

Chitin (**Figure 2-29**) is a polymer of *N*-acetylglucosamine units. The linkage between units is similar to that in cellulose; the main structural difference is that the C2 carbon now bears an amino group that is acetylated, instead of a hydroxyl group. Like cellulose, chitin is used for structural purposes; it forms the hard exoskeleton of arthropods (e.g., insects, spiders, and crabs).

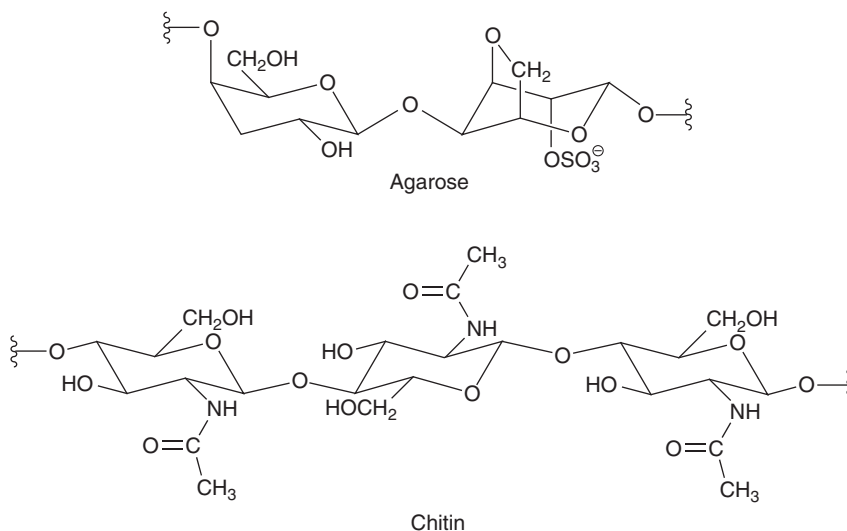


Figure 2-29 Agarose and chitin. Note the unusual ether bridge on the sulfated anhydrogalactose unit in agarose.

Biosynthesis of Proteoglycans

The biosynthesis of mammalian proteoglycans and the associated GAGs starts with the attachment of a xylose residue to the side chain of a serine on the core protein; two galactose residues and a glucuronic acid residue are then attached to the xylose (**Figure 2-30**). The pathways leading to the different GAGs diverge at this point, with different enzymes catalyzing the joining of the sugar residues and more enzymes modifying the residues after they are joined. Chain initiation takes place in the endoplasmic reticulum, and chain elongation and modification occur in the Golgi complex. The proteoglycans move to the cell surface and to the extracellular matrix, carried by vesicles that bud off the Golgi complex. This biosynthetic path is not applicable to hyaluronic acid, however, because this GAG is not attached to a protein but is instead extruded directly to the extracellular matrix.

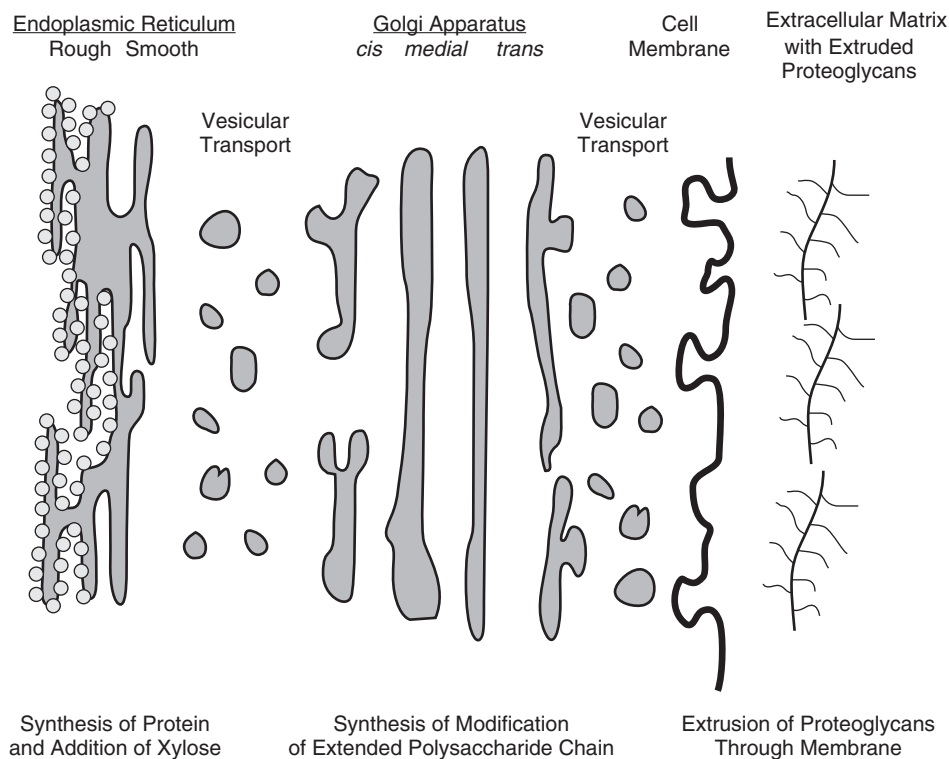


Figure 2-30 Biosynthesis pathway for mammalian proteoglycans and associated GAGs.

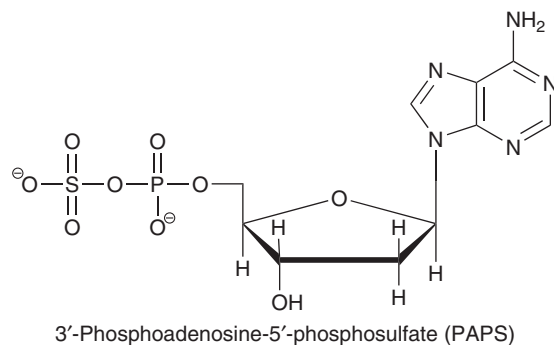


Figure 2-31 Structure of PAPS.

Sulfation reactions take place in both the endoplasmic reticulum and the Golgi complex. The donor of the sulfate group is a compound called PAPS, which is derived from ATP and sulfate (**Figure 2-31**). Sulfation reactions involving PAPS are also important in drug metabolism; the attachment of a sulfate improves the aqueous solubility of a drug or its metabolites, allowing for better excretion.

Other Glycoconjugates

As noted earlier, sugar molecules are attached to a wide variety of other biomolecules, playing roles in molecular recognition processes and stabilizing and perhaps protecting these other biomolecules. This section describes some of the most important classes of biomolecules where sugars are attached.

Glycolipids

Glycolipids, which consist of sugars attached to lipids, are found in biomembranes. Details of their structures are presented in Chapter 5. Sphingolipids are a general class of lipids that include a subclass (*glycosphingolipids*) in which sugars are attached to a ceramide molecule (**Figure 2-32**). A single sugar may be attached, or there may be longer chains, some with

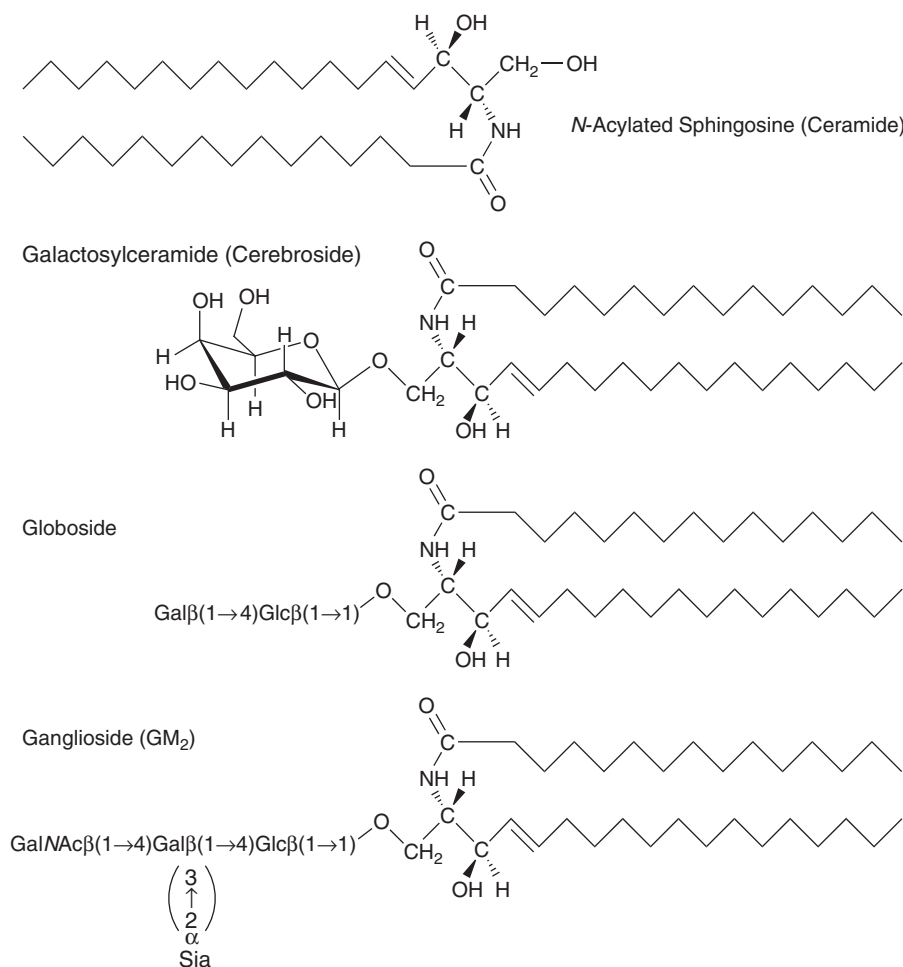


Figure 2-32 Glycosphingolipids. The abbreviation Sia is used to indicate a sialic acid, here *N*-acetylneuraminic acid.

branches. The combination of ceramide and a single sugar residue is called a cerebroside; a globoside is ceramide with a chain of a couple of sugars, and a ganglioside is ceramide with a rather complex sugar chain attached. Gangliosides are prevalent in the membranes of neurons. The sugar moiety or moieties of a glycolipid project outward from the surface of the cell's membrane, and these groups are frequently involved in cell–cell recognition.

Hexosidases are enzymes that break down cerebrosides and gangliosides, removing one sugar residue at a time. These enzymes are found in the lysosomal compartments of the cell. Defects in these enzymes can lead to serious pathologies (lipidoses), such as Tay-Sachs disease. **Table 2-2** lists several of these glycolipid storage diseases and identifies the enzyme deficiency involved in each.

Glycoproteins

Many cellular proteins are modified through the attachment of one or more carbohydrate chains; these are *glycoproteins*. Unlike proteoglycans, glycoproteins are mostly protein, with only a small fraction of their mass being carbohydrate. Also, the attached carbohydrate chains are generally shorter, contain more branching, and have more diversity in sugar sequence compared to the chains found in proteoglycans. In mammals, most cell-surface proteins are glycosylated. Glycoproteins may also be found as secreted proteins (in hormones such as thyroid-stimulating hormone and erythropoietin, in antibodies, and in lactalbumin). Soluble proteins in the cell may also be glycosylated; for example, ribonuclease B differs from ribonuclease A only in the attachment of a sugar chain to a particular asparagine on this enzyme.

Following protein synthesis, polypeptide chains in eukaryotes enter the lumen of the endoplasmic reticulum and later pass through the Golgi complex. During this passage, sugars may be attached to a nitrogen in the side chain of asparagine residues on the polypeptide (so-called *N*-linked carbohydrates) or to oxygens in the side chains of serine or threonine residues (so-called *O*-linked carbohydrates); see **Figure 2-33**. *N*-links are formed in the endoplasmic reticulum and in the Golgi complex, while *O*-links are formed only in the Golgi complex. Complex carbohydrate chains destined to be attached to asparagine are first assembled using a special lipid, dolichol phosphate, embedded in the membrane of the endoplasmic reticulum (**Figure 2-34**). The assembled chain is then transferred to the target protein. Simpler carbohydrate chains start with attachment of a single sugar residue to the target protein, and the chains are elongated by specialized enzymes. After processing in the endoplasmic reticulum, the glycoproteins may be further modified in the Golgi complex. Finally, these glycoproteins are sorted in the Golgi complex, to direct the proteins to their proper cellular locations—for example, the plasma membrane or internal organelles such as lysosomes.

Disease	Lipid Accumulated	Main Organs Affected	Enzyme Deficiency
Fabry	Ceramide trihexoside	Kidney	α -D-galactosidase
Gaucher	Glucocerebroside	Brain, liver, spleen	Glucosylceramide β -D-glucosidase
Krabbe	Galactocerebroside	Brain	Galactosylceramide β -D-galactosidase
Tay-Sachs	Ganglioside GM ₂	Brain	β -D-Hexosaminidase A

Table 2-2 Glycolipid Storage Diseases

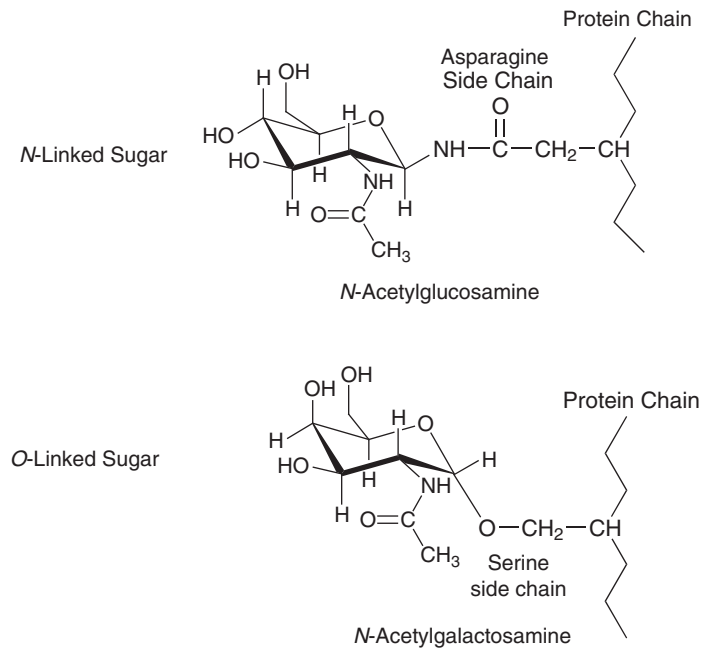


Figure 2-33 O- and N-linked sugars in proteoglycans.

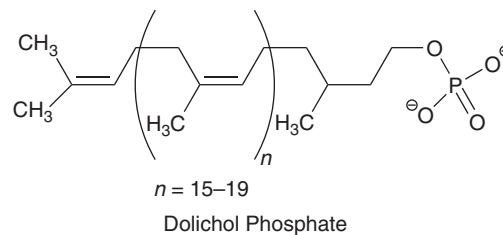


Figure 2-34 Dolichol phosphate.

Lectins

Lectins are proteins that specifically bind carbohydrate chains on the outside of cells to make particular cell-to-cell contacts. One class of lectins, C-type lectins, requires calcium ions to help form protein-carbohydrate complexes. Bacteria use lectins to adhere to epithelial cells in the gut, binding to oligosaccharides on the cell surface. Viruses enter cells by using the protein hemagglutinin to bind to sialic acid residues on glycoproteins embedded in the cell membrane. The “H” part of the designation of strains of influenza virus (e.g., the H1N1 strain of the influenza virus, commonly referred to as swine flu) refers to hemagglutinin; several different, but closely related forms of this viral protein exist. The “N” part of the nomenclature for the influenza virus refers to neuraminidase, another viral protein of which there are several variants. Neuraminidase is an enzyme that cleaves the glycosidic bond joining the sialic residue to the embedded protein after virus enters cell—an action that frees the virus for unpackaging and replication.

Clinical Applications

Clotting and Heparan/Heparin

Heparin is usually prepared from intestinal mucosa, which is rich in heparan sulfate proteoglycans. Heparin is a fraction of heparan. Its structure is that of a repeating disaccharide of glucosamine and iduronate, heavily sulfated. A pentameric sequence within this glycosaminoglycan has great affinity for antithrombin III, a plasma protein that inhibits proteases involved in forming blood clots. This pentasaccharide contains a rare 3-*O*-sulfated glucosamine (**Figure 2-35**). Complexation of heparin with antithrombin enhances the activity of antithrombin greatly, so heparin—and the pentasaccharide in particular—is used in anticoagulant therapy.

Blood Group Antigens

The membranes of red blood cells contain glycolipids that carry oligosaccharides recognized by the immune system. The lipid is a derivative of sphingosine (see Chapters 5 and 13 for more details on glycolipids). The oligosaccharide is composed of a glucose linked directly to the lipid, followed by a galactose residue, an *N*-acetylglucosamine residue, and then a galactose residue. There is usually a terminal *L*-fucose residue in a 1 → 2 linkage. The exact sequence of sugars, and the manner in which they are linked and branched by glycosyltransferase enzymes, determines the blood group type (**Figure 2-36**). One type of glycosyltransferase attaches a branching *N*-acetylgalactosamine to the galactose residue toward the end of the chain, giving the “A” blood type. Another type attaches galactose instead, giving the “B” blood type. In the “O” blood type, neither enzyme is active, and the chain is not branched. A rare blood type is produced by a deficiency in the enzyme that attaches the *L*-fucose; this is the “I” blood type. Other rare blood types are caused by linking the fucose in a 1 → 4 manner rather than a 1 → 2 linkage.

The immune system can produce antibodies against these various short oligosaccharides. People with A-type blood will have antibodies against the B-type oligosaccharides; conversely, people with B-type blood will produce antibodies against the A-type antigen. The antibodies can cause blood cells to clump together, leading to serious circulatory problems. For this reason, it is essential in blood transfusions to match the type of the blood donor to that of the blood recipient. People with the O-type antigens are called universal donors, in recognition of the fact that people with A- or B-type blood lack antibodies to the O antigen. Note, however, that people with O-type blood must receive that type in a transfusion, and not A- or B-type blood, because the A and B antigens would be recognized and attacked by the host immune system.

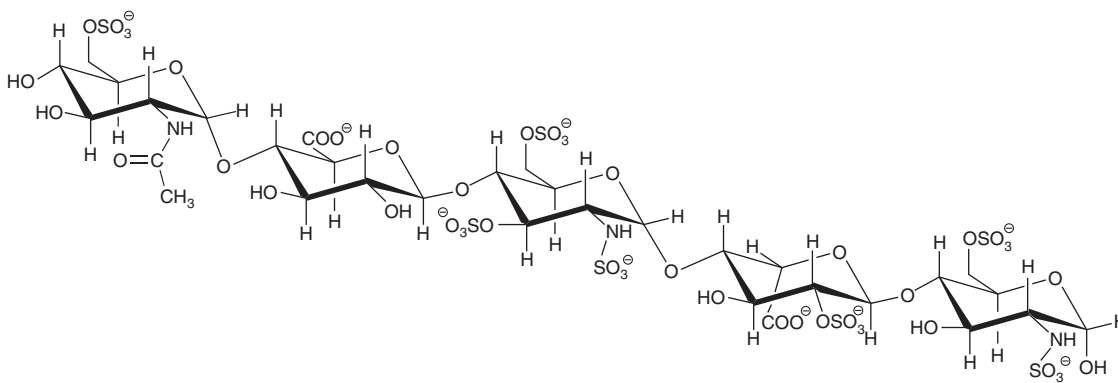


Figure 2-35 Heparin pentasaccharide.

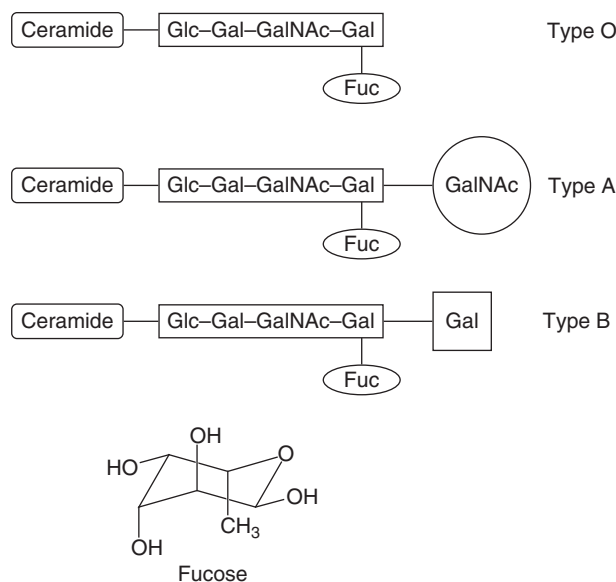


Figure 2-36 Glycolipids for the A-B-O blood types and the structure of fucose, an unusual sugar in these glycolipids.

Blood Glucose Levels, Hemoglobin Glycosylation, and Diabetes

The reducing sugar glucose has a free aldehyde when it is in the linear (not ring) form; this aldehyde is reactive with amino groups, so glucose can be covalently and nonenzymatically linked to a protein through the side-chain amino group of a lysine residue or the terminal —NH_2 of the protein chain. The extent of glycosylation of a common protein, such as hemoglobin (Hb) in circulating erythrocytes, depends on the concentration of free glucose. Glucose levels in blood are unusually high in the disease diabetes mellitus, so that Hb becomes glycosylated. Glycosylated Hb is denoted as HbA_1 ; the major sugar derivative comes from the reaction of glucose with the amino terminal group on the β subunit of Hb. This product, which is denoted as HbA_{1c} , can be detected and quantified using suitable clinical assays. The lifetime of a typical erythrocyte is 110 to 120 days, so the amount of glycosylated Hb is a record of the average level of glucose in the blood over the previous couple of months. A high level of HbA_{1c} indicates high blood glucose levels, signifying that blood glucose is not being well controlled; conversely, one check for satisfactory control of blood glucose in diabetes mellitus is the reduction in detected HbA_{1c} .

Action of Penicillin

A major component of bacterial cell walls is peptidoglycan, which is made of short peptide chains joined to a long heteropolysaccharide of alternating *N*-acetylglucosamine and *N*-acetylmuramic acid residues. The peptide chains are unusual in that they often contain a large proportion of *D*-amino acids, instead of the usual *L*-amino acids. These short peptide chains act as cross-links between polysaccharide chains. The cross-linking reaction is catalyzed by peptidoglycan transpeptidase. It is this enzyme that is the target of the antibiotic penicillin. Without a strong, intact cell wall, the bacteria cannot survive, and penicillin weakens the cell wall by inhibiting the cross-linking reaction.

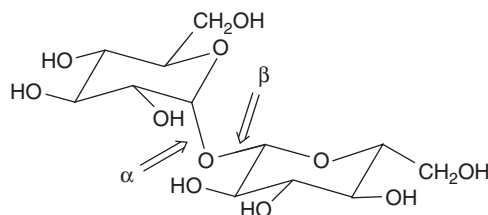


Figure 2-37 α , β -Trehalose. Arrows point to the α and β linkages.

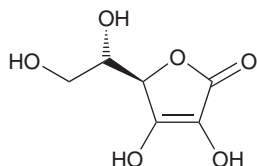


Figure 2-38 Vitamin C.

QUESTIONS FOR DISCUSSION

1. Is glycerol a carbohydrate? Is it a monosaccharide? What about acetone?
2. **Figure 2-37** gives the structure of trehalose. Predict the products of hydrolysis of this sugar. Is trehalose a reducing sugar? What about the products?
3. In solution, fructose exists about 60% in the pyranose form; of this 60%, 57% is the β -D-fructopyranose, and only about 3% is the α -D-fructopyranose. Why is the β anomer favored?
4. Suggest a structure for the product of the reaction of glucose with a primary amine (as in the N-terminal valine in the beta subunit of hemoglobin).
5. **Figure 2-38** presents a structure for ascorbic acid, vitamin C. Is vitamin C a carbohydrate?
6. Why does reduction of D-fructose give a mixture of D-sorbitol and D-mannitol? (See Figure 2-12.)
7. Heparin preparations have reportedly been contaminated with chondroitin sulfate (CS). Such contamination is quite serious, as it can lead to severe allergic reactions. The contaminating CS was oversulfated (to a degree not found naturally), and the presence of this contaminant was likely the result of deliberate adulteration by the supplier of the heparin. Compare the structures of heparin and CS, and comment on how CS could pass as heparin in cursory tests of the properties of the heparin preparation.

REFERENCES

- C. G. Gahmberg and M. Tolvanen. (1996). "Why mammalian cell surface proteins are glycoproteins," *Trends Biochem. Sci.* 21:308–311.
- R. T. Morrison and R. N. Boyd. (1992). *Organic Chemistry*, 6th ed., Prentice-Hall, Englewood Cliffs, NJ.
- M. Petitou, B. Casu, and U. Lindahl. (2003). "1976–1983, a critical period in the history of heparin: The discovery of the antithrombin binding site," *Biochimie* 85:83–89.

- N. Sharon and H. Lis. (2004). "History of lectins: From hemagglutinins to biological recognition molecules," *Glycobiol.* 14:53R–62R.
- K. R. Taylor and R. L. Gallo. (2006). "Glycosaminoglycans and their proteoglycans: Host-associated molecular patterns for initiation and modulation of inflammation," *FASEB J.* 20:9–22.
- R. R. Yocum, D. J. Waxman, and J. L. Strominger. (1980). "Interaction of penicillin with its receptors in bacterial membranes," *Trends Biochem. Sci.* 5:97–101.

