

# Carbohydrates

Robert J Sturgeon, *Heriot-Watt University, Edinburgh, UK*

The carbohydrates comprise one of the major groups of naturally occurring organic molecules and are amongst the most abundant constituents of plants, animals and microorganisms. In general, carbohydrates are polyhydroxy-aldehydes or-ketones. They may contain, in addition, amino, acetamido and carboxyl functional groups.

## Introduction

The term carbohydrate includes monosaccharides, oligosaccharides and polysaccharides. Also included are substances derived from monosaccharides such as alditols, which are derived by reduction of the carbonyl group and carboxylic acids, which are derived by oxidation of one or more terminal groups. Replacement of a hydroxyl group with a hydrogen atom produces a deoxy-sugar and replacement of a hydroxyl group with an amino group produces an amino sugar. The term 'sugar' is frequently applied to monosaccharides and lower molecular weight oligosaccharides.

## Classification

Carbohydrates are usually classified in three groups: monosaccharides, oligosaccharides and polysaccharides. Monosaccharides are simple sugars that cannot be hydrolysed to smaller molecules. They exist in nature in the free form or linked by glycosidic bonds to other monosaccharides in the formation of oligosaccharides or polysaccharides. Oligosaccharides are defined as simple polymers of monosaccharides containing between two and approximately 10 monosaccharide residues. They are termed disaccharides, trisaccharides, tetrasaccharides, pentasaccharides, and so on, according to the number of monosaccharide units they contain. Polysaccharides (glycans) are higher molecular weight polymers of monosaccharides. Glycoconjugates are defined as either glycoproteins, glycolipids and proteoglycans. Glycoproteins are conjugated proteins containing either oligosaccharide groups or polysaccharide groups, having a fairly low molecular mass. Glycolipids are conjugated lipids containing oligosaccharide groups, and proteoglycans are proteins linked to polysaccharides of high molecular mass.

## Monosaccharides

Monosaccharides are polyhydroxyaldehydes  $\text{H}[\text{CHOH}]_x\text{-CHO}$  or polyhydroxyketones  $\text{H}[\text{CHOH}]_x\text{-CO}[\text{CHOH}]_y\text{-H}$  with three or more carbon atoms. Monosaccharides bearing an aldehydic carbonyl group are called aldoses, whereas those with a ketonic carbonyl group are called

## Introductory article

### Article Contents

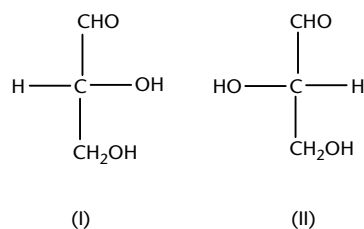
- Introduction
- Monosaccharides
- Higher-order Structures
- Naturally Occurring Oligosaccharides and Glycosides
- Plant and Algal Polysaccharides
- Polysaccharides in Fungi and Invertebrates
- Polysaccharides and Glycoconjugates of Bacteria
- Polysaccharides and Glycoconjugates in Higher Animals
- Carbohydrates in Recognition Processes
- Summary

ketoses. As monosaccharides are normally found as cyclic structures, containing either a hemiacetal or hemiketal group arising from ring closure of the linear polyhydroxycarbonyl compounds, monosaccharides are considered to contain a potential aldehydic carbonyl group (hemiacetal) or potential ketonic carbohydrate group (hemiketal). Cyclic hemiacetals or hemiketals of sugars with a five-membered (tetrahydrofuran) ring are described as furanoses and those with a six-membered (tetrahydropyran) ring are called pyranoses.

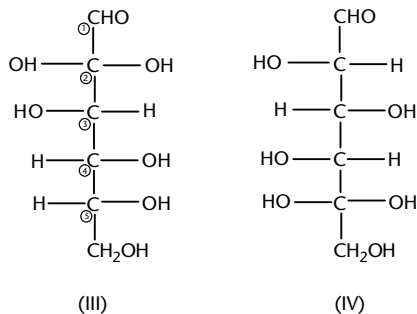
Monosaccharides are classified, according to the number of carbon atoms they contain, as trioses, tetraoses, pentoses, hexoses, etc.

## Stereoisomerism and configuration

In the early 1890s Emil Fischer published detailed studies on the configuration of aldoses. Molecules containing one centre of chirality (asymmetrical carbon atom – which is a carbon atom to which four different atoms or groups are attached) exist in two isomeric forms. They are stereoisomers, differing only in the arrangement of groups in space. Thus the simplest sugar, glyceraldehyde, exists in two nonsuperimposable stereoisomeric molecules, which are mirror images. The two stereoisomers are called enantiomers, or an enantiomeric pair. It is known that these projections correspond to absolute configurations, which are referred to as D-glyceraldehyde (I) and L-glyceraldehyde (II).



The carbon atoms of a monosaccharide are numbered consecutively, with the (potential) aldehyde group as number 1. The monosaccharides are then assigned to the D or L series according to the highest-numbered centre of chirality. The structure below is the Fischer projection for the acyclic form of D-glucose (III) and L-glucose (IV).



Comparison of these structures with those of D- and L-glyceraldehyde illustrate the relationship of the configurations. The structures of monosaccharides (triose, tetrose, pentose and hexose) in the aldehydic acyclic form together with their trivial names and three-letter abbreviations for hexoses and pentoses are given in **Figure 1**. Only the D forms are shown.

## Ring structure

Most monosaccharides exist as cyclic hemiacetals or hemiketals, and are customarily presented in the Haworth representation. Using D-glucose as an example, the relationship between the acyclic Fischer representation and the Haworth representation is shown in **Figure 2**.

Groups appearing to the right of the Fischer projection appear below the plane of the ring in the Haworth representation, and those on the left of the Fischer projection appear above the plane of the ring in the Haworth representation. The formation of the ring form produces a new asymmetric carbon atom at C-1 in aldoses (or C-2 in ketoses). This is called the anomeric carbon atom and the isomers are called anomers, distinguished by the terms  $\alpha$  and  $\beta$ .

In the Haworth structures, one is visualizing the groups attached to carbon atoms as being either above or below the plane of the ring, implying that it is a flat rigid ring. However, monosaccharides assume conformations that are not planar. For consideration of the conformation of monosaccharides (the arrangement in space of the atoms),  $\beta$ -D-glucopyranose assumes a chair conformation. For clarity, in most representations the hydrogen atoms bonded to carbon atoms are omitted.

## Physical and chemical properties

Monosaccharides are mainly crystalline solids, soluble in water and capable of forming viscous solutions. The

specific rotation ( $\alpha$ ) of freshly prepared aqueous solutions of monosaccharides such as glucose change rapidly, before reaching a constant value in a phenomenon known as mutarotation. Mutarotation generally follows first-order kinetics, in which ring opening of a hemiacetal or hemiketal sugar takes place followed by ring closure with the production of quantities of the other anomer.

A few sugars (L-arabinose, D-galactose and D-fructose) exhibit a more complex form of mutarotation, where at equilibrium a mixture of both pyranose and furanose forms is obtained.

As most monosaccharides contain a primary hydroxyl group, several secondary hydroxyl groups and a potential carbonyl group, they can undergo a variety of chemical reactions to produce ethers, acetals, esters and glycosides (**Figure 3**). Methyl ethers (**Figure 3(i)**) are among the most common derivatives and have been widely used in analytical work, mainly in the identification of the linkage positions between individual monosaccharides in oligosaccharides and polysaccharides. Benzyl ethers are amongst the most commonly used protecting groups in carbohydrate chemistry, where they are used mainly for the temporary substitution of hydroxy groups on sugar derivatives required for synthetic purposes. Provided that the monosaccharides do not contain base-sensitive groups, traditional methods, involving reagents as benzyl halides with sodium hydroxide or silver oxide may be used. Benzyl ethers have similar stability to methyl ethers. Debenzylation is easily achieved by catalytic hydrogenation.

Glycosides (**Figure 3(ii)**) are sugar derivatives bearing a substituent (the aglycone) at the anomeric oxygen atom of an aldose or ketose to produce a hemiketal. These products are stable to alkali, but labile to acid, do not undergo reversible ring opening, and therefore do not undergo mutarotation, as do the parent aldoses or ketoses.

The conversion of aldoses to glycosides is a common procedure for the protection of the anomeric centre of the sugar, prior to the synthesis of other derivatives. Frequently, sugars are converted to glycosides by alcoholysis in the presence of an acid catalyst, in a reaction known as Fischer glycosidation. The method is nonspecific, and mixtures of  $\alpha$ - and  $\beta$ -pyranosides are frequently formed.

Acetals (**Figure 3(iii)**) are produced by condensation of two alcoholic groups and a carbonyl group. A wide variety of derivatives may be produced due to the possibility of intermolecular and intramolecular reactions. The products are often crystalline and serve as convenient intermediates in synthesis. Reaction, for example, of glucose with acetone and in the presence of a dehydrating agent produces 1,2:3,4-di-O-isopropylidene glucose.

The hydroxyl groups of sugars and their derivatives may be readily esterified (**Figure 3(iv)**), with acetates and benzoates being used frequently in synthetic work. Glucose treated with acetic anhydride in the presence of sodium acetate will furnish mainly penta-O-acetyl- $\beta$ -D-glucopyranose. When sodium acetate is replaced by zinc chloride,

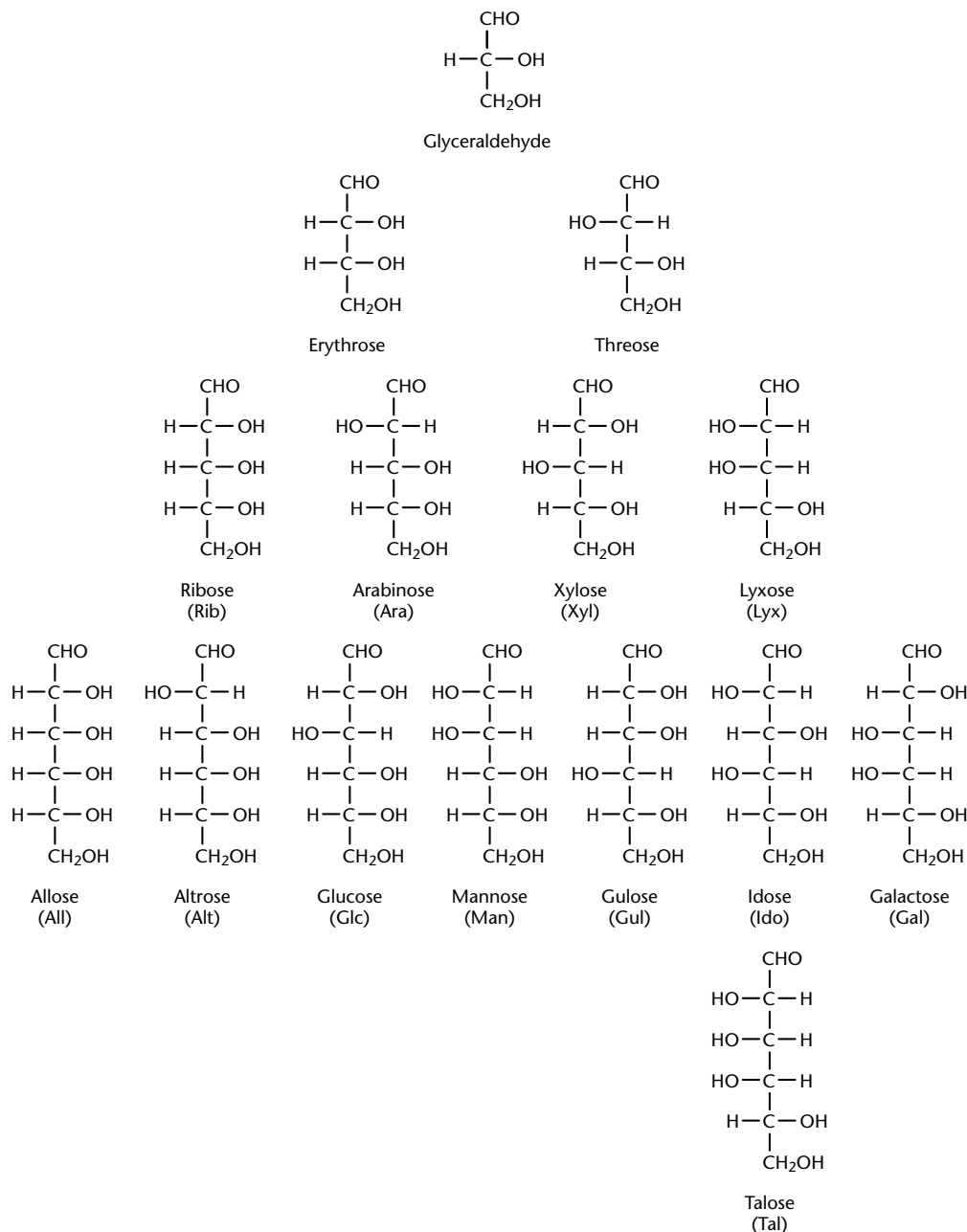


Figure 1 Aldose sugars of the D series.

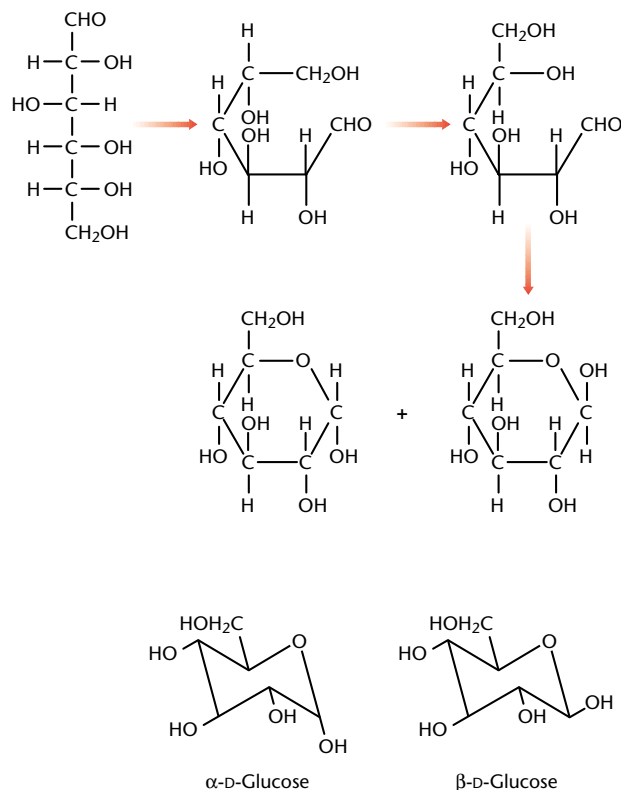
the main product is penta-*O*-acetyl- $\alpha$ -D-glucopyranose. Sugar esters are readily cleaved under acidic or basic conditions.

Monosaccharides are readily reduced to alditols (Figure 3(v)), using water-soluble reductants such as sodium borohydride. Reduction of the aldehyde or potential aldehyde group in aldoses gives one product. Glucose is converted to glucitol, but ketoses such as fructose gives two

products, glucitol and mannitol, due to the creation of a new asymmetrical carbon atom at C-2.

Mild oxidizing reagents such as bromine water or sodium hypochlorite oxidize aldoses at C-1 to produce the corresponding aldonic acid (Figure 3(vi)). Glucose is converted to gluconic acid, which readily gives a stable lactone. Ketoses are not oxidized by means of these procedures.





**Figure 2** Acyclic (Fischer representation) and cyclic (Haworth representation) forms of D-glucose.

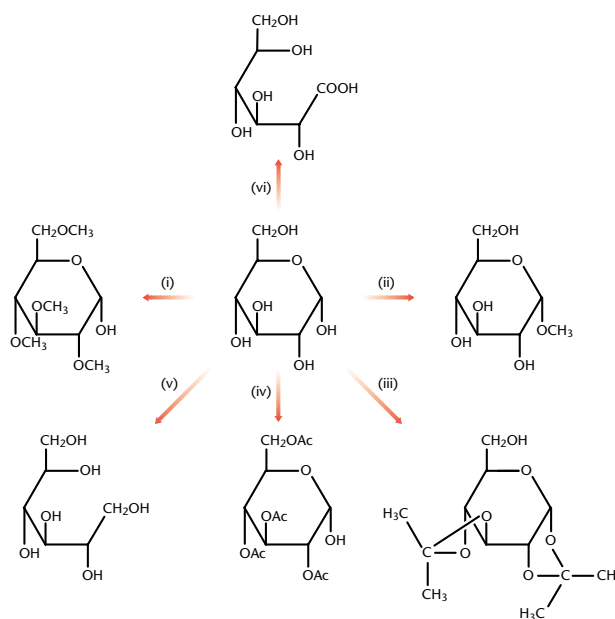
from dextro to laevo owing to the high laevorotation of the D-fructose.

In human metabolism, sucrose, as part of the diet, is readily converted by  $\alpha$ -glucosidases present on the intestinal wall. The resulting D-glucose and D-fructose are readily transported through the intestinal wall into the bloodstream, and transported to and utilized by specific cells in the body.

A number of other oligosaccharides based on the structure of sucrose are important in the metabolism of plants. For example, raffinose, a naturally occurring plant trisaccharide, has a distribution in nature almost as ubiquitous as that of sucrose, but in much smaller amounts (approximately 0.05% in sugar beet and 1.9% in soy beans). Raffinose is an *O*- $\alpha$ -D-galactopyranosyl-(1 $\rightarrow$ 6)- $\alpha$ -D-glucopyranosyl-(1 $\rightarrow$ 2)- $\beta$ -D-fructofuranoside. It is readily converted by the action of  $\beta$ -fructofuranosidase (invertase) to the corresponding disaccharide melibiose (*O*- $\alpha$ -D-galactopyranosyl-(1 $\rightarrow$ 6)- $\alpha$ -D-glucopyranose).

### Lactose

Lactose (VI) occurs in mammalian milk, from between 2.0 and 8.5 per cent, depending on the mammal. It was



**Figure 3** Derivatization of D-glucose. (i) Ether formation to produce 2,3,4,6-tetra-O-methyl-D-glucose. (ii) Glycoside formation to produce methyl- $\alpha$ -D-glucoside. (iii) Acetalation to form 1,2,3,4-di-isopropylidene D-glucose. (iv) Ester formation to form 2,3,4,6-tetra-O-acetyl-D-glucose. (v) Reduction to form D-glucitol. (vi) Oxidation to form D-glucuronic acid.

originally named 'milk sugar' and is 4-*O*- $\beta$ -D-galactopyranosyl-D-glucose, and thus a reducing sugar. Lactose is prepared industrially from whey, which is obtained as a byproduct in the manufacture of cheese.

The disaccharide maltose (V) (*O*- $\alpha$ -D-glucopyranosyl-(1 $\rightarrow$ 4)-D-glucose) is an intermediate formed during the metabolic mobilization of starch, a common reserve polysaccharide in plant tissues. It is produced, together with other malto-oligosaccharides, in the saccharification of starch by amylases during the development, maturation and germination of cereal grains. The major pathway for the metabolic utilization of maltose in plants is its cleavage to free glucose, by the action of  $\alpha$ -glucosidase (maltase), an enzyme found in starch-producing tissues.

Various food preparations contain large amounts of maltose, which when ingested by humans is hydrolysed to glucose by an  $\alpha$ -glucosidase in the intestinal tract before transport across the intestinal membrane into the bloodstream can take place. The glucose is then available for distribution to specific cells for further metabolism. Starch itself, which forms part of the human diet, is first broken down by salivary and pancreatic amylases to maltose and malto-oligosaccharides prior to conversion to glucose by the intestinal  $\alpha$ -glucosidases.

Most other oligosaccharides, which may occur freely, but in very small concentrations, may be prepared either by

controlled partial acid hydrolysis or by enzymic hydrolysis of the native polysaccharide. These methods usually produce mixtures (mono-, di-, tri-saccharides, etc.). Individual saccharides may then be obtained following some chromatographic procedure.

## Glycosides

A host of natural compounds in which sugars are glycosidically linked to nonsugar constituents (aglycone) have been identified. These include simple alcohols, alkaloids, carotenes, natural pigments and steroids. Many of these compounds have been shown to be pharmacologically active, including cardiac glycosides and antibiotics such as streptomycin and the anthracyclins. D-Glucose,  $\beta$ -linked to the aglycone, is commonly but not exclusively found, and many plant glycosides have been shown to contain disaccharide and occasionally trisaccharide units. Microbial products have revealed a much larger range of glycosidically linked sugars containing structural variations such as amino, carboxyl and deoxy groups.

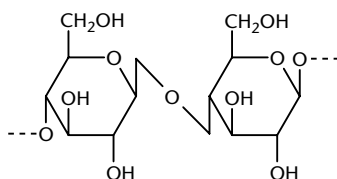
## Plant and Algal Polysaccharides

Plants can be considered in two separate groups: land plants and aquatic plants (algae).

### Plant polysaccharides

#### Cellulose

Land plants generally contain high levels of polysaccharides, with cellulose (VIII), a (1 $\rightarrow$ 4)- $\beta$ -D-glucan and the primary component of the cell wall, being the most abundant. Associated with cellulose are other heteropolysaccharides such as arabinoxylyans, arabinogalactans and pectic substances.



(VIII)

Reserve, or storage, polysaccharides are synthesized by plant cells to be later enzymically hydrolysed to produce the monosaccharides, which re-enter the cell's metabolism. These polysaccharides are generally synthesized during periods of high photosynthetic activity. Starch, which is economically the most important seed storage polysaccharide, is formed inside the plastids. Resting seeds of many plant species may contain little or no starch. In these

cases, polysaccharide reserves are stored outside the plasmalemma and are described as cell wall storage polysaccharides, represented by mannans, xyloglucans and galactans. Cellulose is the most abundant organic substance found on Earth. It occurs as 94% of the cotton fibre, approximately 80% in flax, 60–70% in jute, 40–50% in wood and as little as 4% in endosperm cells.

In cellulose, the molecules are entirely linear, consisting of up to 14 000 D-glucose units linked uniformly by (1 $\rightarrow$ 4)- $\beta$ -glycosidic bonds. The  $\beta$ -configuration between glucose units, together with hydrogen bonding which arises from the ring oxygen of one glucose residue and the C-3 hydroxyl group on the neighbouring glucose residue, results in the formation of an extended ribbon-like chain. In the plant cell, these polymer chains stack on top of each other, joined by hydrogen bonds, in a way that staggers the chains – as bricks are staggered to give stability to a wall. This results in the cellulose existing in the form of microfibrils, where both crystalline and amorphous areas can be observed. Native cellulose is not very reactive chemically, except in the amorphous regions. Strong acids, such as hydrochloric or sulfuric acid, are necessary to cleave glycosidic bonds to produce glucose and oligosaccharides (cellodextrins).

The traditional uses of wood cellulose in the paper industry and cotton cellulose in the textile industry have been extended to the industrial production of a number of derivatives of cellulose. When cellulose is preswollen by the use of alkali, the polymer chains are much more susceptible to chemical modifications. Products such as cellulose acetate can be used to produce fibres (rayon), lacquers and films. Cellulose nitrate is used in the plastics, film and explosive industries. A water-soluble, viscous derivative of cellulose, carboxymethyl cellulose, is used in the food industry as a thickening agent in adhesives, laxatives and low-calorie foods.

#### Starch

Starch is the principal food reserve polysaccharide in the plant kingdom, and occurs within the protoplasts of many plant cells in the form of granules. These granules contain two separately different polysaccharides, amylose (IX) and amylopectin (X), in the approximate ratio of 1 : 4. The former consists of long, mostly linear, chains of (1 $\rightarrow$ 4)- $\alpha$ -linked D-glucose residues. Depending on the source of the starch, amylose may contain several hundred to several thousand glucose residues. The branching is due to the presence of (1 $\rightarrow$ 6)- $\alpha$ -D-glucosidic linkages.

O- = (1 $\rightarrow$ 4)- $\alpha$ -D-glucosyl-

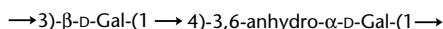
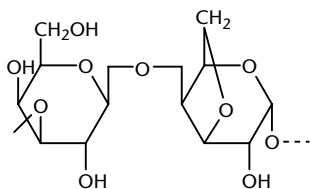
O $\rightarrow$  = (1 $\rightarrow$ 6)- $\alpha$ -D-glucosyl-

Amylopectin has a branched structure, in which the linear chains of approximately 20 to 25 (1 $\rightarrow$ 4)- $\alpha$ -D-glucose residues are linked via (1 $\rightarrow$ 6)- $\alpha$ -D-glucosidic linkages to form a macromolecule of molecular weight in the range  $10^7$  to  $10^8$ , the variation being dependent on the botanical



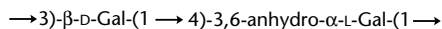
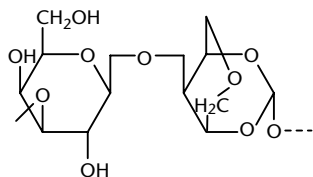
hold water, form gels and stabilize emulsions has led to numerous industrial and food applications.

The major structural polysaccharides of red seaweed are the carrageenans, a family of partially sulfated linear galactans. The polysaccharide (XII) has a regular disaccharide repeat structure of 3-linked  $\beta$ -D-galactose and 4-linked 3,6-anhydro- $\alpha$ -D-galactose residues. Sulfate ester groups are located at different sites on the polysaccharide to produce three different molecules: iota ( $\iota$ ), kappa ( $\kappa$ ) and lambda ( $\lambda$ ) carrageenans. Carrageenan produces high-viscosity solutions and gels in water, and is widely used in the food industry. It reacts with the milk protein, casein, and is used to prepare high-strength milk gels.



(XII)

Red seaweeds are the source of agar, which is a mixture of at least three polysaccharides: agarose (XIII), a neutral polysaccharide with alternating (1 $\rightarrow$ 3)- $\beta$ -D-galactosyl and (1 $\rightarrow$ 4)-linked 3,6-anhydro- $\alpha$ -L-galactosyl residues, and two other acidic polysaccharides.



(XIII)

Agarose is the gelling portion of agar, and has a double-helical structure. The double helices aggregate to give a three-dimensional framework, capable of holding water molecules within the interstices of the framework. As these gels are thermoreversible, agar is an ideal material for the production of microbiological media for culturing microorganisms.

## Polysaccharides in Fungi and Invertebrates

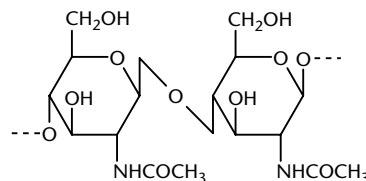
### Fungal polysaccharides

The cell walls of fungi are rich in polysaccharides, generally containing an inner layer composed of either chitin or cellulose, embedded in other polymers, and an outer layer. The outer layer is generally soluble in dilute alkali, with the inner layer being an insoluble residue. The inner layer of most taxonomic groups normally contains two different polysaccharides: a highly branched  $\beta$ -glucan and chitin, which is a polymer of (1 $\rightarrow$ 4)-linked  $\beta$ -N-acetylglucosamine residues.

*Saccharomyces cerevisiae* is probably the yeast studied in greatest detail, and principal components of its cell wall are polysaccharides, glucan and mannan, which together account for approximately 80–90% of the dry weight of the cell. The glucan component is responsible for the tensile strength, rigidity and shape of the cell. In addition to the cell wall glucan, *S. cerevisiae* contains an intracellular storage polysaccharide, glycogen. This is a highly branched  $\alpha$ -glucan of similar structure to plant amylopectins. When grown in the presence of a suitable carbohydrate source, *Aureobasidium pullulans* (formerly *Pullularia pullulans*) excretes large quantities of a viscous water-soluble  $\alpha$ -glucan. This polysaccharide, referred to as pullulan, is a linear polysaccharide containing mainly maltotriose units connected by (1 $\rightarrow$ 6)- $\alpha$ -linkages. Pullulan is produced on an industrial scale, with applications in forming films for food packaging and in the formulation of certain foods.

### Invertebrate polysaccharides

Chitin (XIV) is the major invertebrate polysaccharide and forms 15–30% of the shells of crustacea, the rest being mainly calcium carbonate. It is a fibrillar (1 $\rightarrow$ 4)-linked 2-acetamido-2-deoxy- $\beta$ -D-glucan. It can be considered to have a similar structure to cellulose, with the D-glucose units in that polysaccharide being replaced by 2-acetamido-2-deoxy-D-glucose (*N*-acetylglucosamine) units. Thus the two polysaccharides have similar physical properties. The chains of chitin are of the extended ribbon type, possessing intramolecular hydrogen bonds.



(XIV)



Although native chitin is extremely resistant to chemical attack, it can be dissolved in a number of hydrogen bond-breaking solvents. Prolonged heating in  $10 \text{ mol L}^{-1}$  hot alkali removes the majority of the acetyl groups to give chitosan, whereas treatment with strong ( $5\text{--}10 \text{ mol L}^{-1}$ ) acid at  $100^\circ\text{C}$  is necessary to hydrolyse chitin to glucosamine and acetic acid.

Chitosan is an important water-soluble viscous polysaccharide. As the majority of *N*-acetyl groups of chitin are replaced by amino groups, at low pH values these amino groups exist in a charged ( $\text{N}^+\text{H}_3$ ) form. At high pH, chitosan becomes insoluble.

Chitosan has proved to be useful in the food and biomedical fields – from the removal of anionic particulates from fruit juices to wound-healing technology.

## Polysaccharides and Glycoconjugates of Bacteria

Many polysaccharides and glycoconjugates are synthesized by bacteria and originate either in the cell wall or as intracellular or extracellular polymers. More than 100 different monosaccharides have been identified as components of bacterial polysaccharides isolated from different species or groups of bacteria. Included among these monosaccharides are mono- and di-deoxy sugars, amino-di-deoxy and amino-uronic acids.

Bacteria are normally classified according to the Gram stain as either Gram-positive or Gram-negative. Most Gram-negative bacteria contain a thin inner sheet of peptidoglycan together with an outer membrane composed of lipoprotein, phospholipid and lipopolysaccharide. In Gram-positive bacteria peptidoglycan is the major cell wall component, together with smaller quantities of teichoic acids (polyol phosphates) or other polysaccharides. The peptidoglycan is the major supporting structure of the bacterial cell wall and is a glycan composed of alternating units of *N*-acetylglucosamine and *N*-acetylmuramic acid. Individual glycan chains are cross-linked with either tetra- or penta-peptides. Lipopolysaccharides, which are located on the outer membrane of Gram-negative bacterial cell walls, are responsible for a number of biological properties including toxicity and immunogenicity. The complete lipopolysaccharide consists of three parts: lipid A, the core region and the *O*-specific side-chains. Lipid A is a phosphorylated polysaccharide containing esterified fatty acid groups, and carries the main endotoxic properties of the lipopolysaccharide. The core region is attached to lipid A, and in turn is attached to the *O*-antigenic side-chains, which contain the structural information recognized by antibodies.

Many bacteria produce extracellular polysaccharides, which together with the *O*-antigenic parts of the lipopolysaccharides form the principal immunogens and antigens of the bacteria.

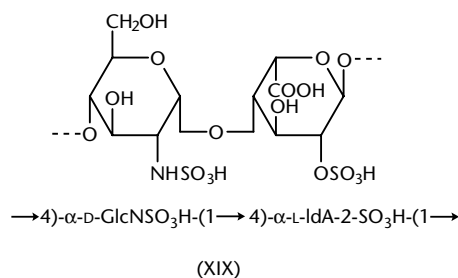
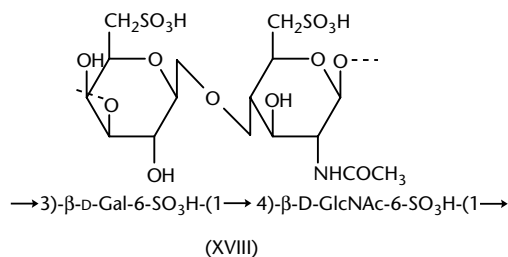
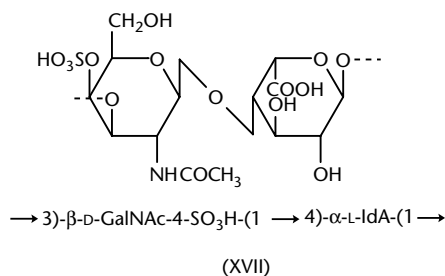
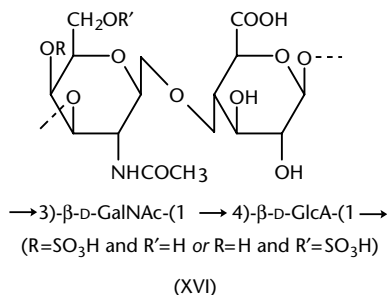
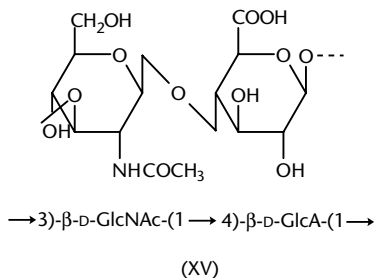
A number of microbial polysaccharides are produced on an industrial scale. Dextran is a high molecular weight  $\alpha$ -D-glucan produced from strains of *Leuconostoc mesenteroides*. This polysaccharide and its derivatives are used in the pharmaceutical and fine chemical industries. Xanthan gum is produced as an extracellular polysaccharide, by *Xanthomonas campestris*. Solutions of the polysaccharide have high viscosity at low concentration. Xanthan is used widely as an additive in the food industry, as well as having many industrial applications.

## Polysaccharides and Glycoconjugates in Higher Animals

### Polysaccharides

The extracellular space of many mammalian tissues contains polysaccharides, composed of amino sugars (glycosaminoglycans) that are covalently attached to a protein backbone. In the older literature they were described as mucopolysaccharide-protein complexes; now they are referred to as proteoglycans. A major function of proteoglycans is in the formation of a structural matrix for protein components of connective tissue and skin. A single proteoglycan may contain more than one type of glycosaminoglycan chain. For example, the cartilage proteoglycan (molecular weight  $2.5 \times 10^6$  Da, 90% carbohydrate) has a core protein with 80 to 100 chondroitin sulfate chains, 30 to 50 keratan sulfate chains, together with oligosaccharide chains that are directly linked to amino acids in the polypeptide backbone.

The glycosaminoglycans are generally composed of repeating disaccharide units, composed of an acidic and a basic monosaccharide. The amino groups of the basic monosaccharides are *N*-acetylated or *N*-sulfated amino groups. *O*-Sulfate ester groups are also present, resulting in the production of highly acidic polysaccharides. Glycoside linkages are alternatively (1→3) and (1→4) linked. Glycosaminoglycans that have been characterized include hyaluronic acid (XV) (composed of *N*-acetylglucosamine and glucuronic acid residues), chondroitin sulfates (XVI) (*N*-acetylgalactosamine 4- or 6-sulfates and glucuronic acid), dermatan sulfate (XVII) (*N*-acetylgalactosamine 4-sulfate and L-iduronic acid) and keratan sulfate (XVIII) (galactose 6-sulfate and *N*-acetylglucosamine 6-sulfate).



Heparin (XIX) is a linear polysaccharide containing mainly D-glucosamine and L-iduronic acid residues, with some D-glucuronic acid residues. The proportion of these

residues is variable, as is the number of sulfate ester groups present. Heparin is a degradation product of a native proteoglycan found exclusively in the granules of mast cells, which are commonly associated alongside blood vessels, and is best known as a powerful anticoagulant. The anticoagulant activity arises from the activation of antithrombin III, a protease inhibitor that is responsible for the inhibition of several proteolytic enzymes of the blood coagulation cascade. Hyaluronic acid is a nonsulfated glycosaminoglycan located in joints, umbilical cord and vitreous humour of the eyes. This high molecular weight viscous polysaccharide is not covalently attached to protein as in proteoglycans, but physically associates with proteoglycan.

Glycogen, a highly branched α-D-glucan with a structure generally similar to that of amylopectin, is the major storage polysaccharide of animals. It is located in the liver and skeletal muscle, and acts as a primary energy store in cells. The breakdown of glycogen requires the presence of two enzymes, a phosphorylase and a debranching enzyme, to produce mainly glucose 1-phosphate and glucose. The glucose 1-phosphate can then be converted to glucose 6-phosphate, then glucose, before use in various metabolic pathways.

## Mammalian glycoconjugates

Glycoconjugates include glycoproteins and glycolipids – proteins or lipids to which carbohydrates are glycosidically attached. Glycolipids are found in animal cells, the major group being the glycosphingolipids. The lipid moiety is ceramide, a molecule in which a long-chain fatty acid is attached to sphingosine, a 2-amino-1,3-dihydroxy-4-*trans*-D-octadecene, via an amide bond. The primary hydroxy group of the ceramide is linked glycosidically to oligosaccharides, usually by a β-D-glucosyl residue. Families of the glycosphingolipids exist, depending on whether a tetra-, tri-, di- or mono-glycosyl saccharide is involved. Gangliosides are a special group of animal glycolipids, in which the glycosylceramides contain at least one sialic acid residue. Sialic acid is a nine-carbon carboxylated sugar, which thus makes the gangliosides anionic.

Glycolipids are generally associated with cell membranes. As a result of their amphipathic nature, the hydrophobic ceramide portion of the molecule forms hydrophobic interactions with the phospholipid of the cell membranes. The covalently attached hydrophilic carbohydrate portion is then located at the outer surface of the cell membrane. A number of human disease states, including Krabbe disease, Gaucher disease and several forms of gangliosidosis, are recognized by the accumulation in the body of specific glycolipids.

Mammalian glycoproteins are found in both soluble and membrane-bound forms. More than 300 proteins have been characterized from human serum or plasma, the vast

majority of these being glycoproteins. A wide variety of properties has been assigned to these glycoproteins, including acting as protease inhibitors, blood clotting factors, hormones, antibodies and the transport of various ions. The presence of carbohydrate in a glycoprotein is known to modify the physicochemical properties of proteins. The presence of oligosaccharide chains will influence the final conformation of proteins. Sialic acid residues, which are normally present at the terminal nonreducing ends of these oligosaccharide chains, are important in determining the correct arrangement of glycoproteins in cell membranes. In a number of cases, the presence of the carbohydrate moiety of glycoproteins confers resistance of these molecules to attack by proteolytic enzymes. Cell-surface carbohydrate moieties may be involved in adhesion, communication recognition, antigenic specificity and regulation of cell growth.

Oligosaccharides are conjugated to mammalian proteins, either by *O*-glycosyl linkages to the hydroxyl groups of serine or threonine (*O*-glycans) or by *N*-glycosyl linkages to the amide nitrogen of asparagine (*N*-glycans). *N*-glycan chains may have two to four branches, and generally contain between 11 and 18 sugar residues, whereas *O*-glycan chains are simpler and may be limited to trisaccharides.

In membrane glycoproteins, the hydrophilic glycan chains on the polypeptide backbone are located outside the cell membrane. A sequence of mainly hydrophobic amino acids is involved in hydrophobic–hydrophobic binding in the membrane itself, with the remainder of the polypeptide chain being in the cytoplasm.

## Carbohydrates in Recognition Processes

Many carbohydrate-specific proteins have been isolated and characterized. Carbohydrates serve as recognition determinants on glycoconjugates. This may involve molecule–molecule, cell–molecule and cell–cell interactions. The most obvious molecule–molecule interactions include enzymic synthesis and degradation of carbohydrates, and the immunological response to carbohydrate antigens.

One of the earliest examples of molecule–cell interactions involved the observation that certain seed proteins have the ability to agglutinate animal blood cells. These proteins are termed lectins. The initial event is the noncovalent binding of a monosaccharide or oligosaccharide segment of a glycoconjugate or polysaccharide to a binding site on the protein. Inhibition of the binding (agglutination) by free monosaccharides or oligosaccharides gives information on the nature of the carbohydrate involved in the binding. Although many lectins have been

isolated from plants, these proteins have also been discovered in animals and microorganisms. These interactions are important in the attachment of toxins, viruses and bacteria to cells.

The adherence of bacteria to mammalian cells is a prerequisite for colonization and infection, while the initial adherence of bacterial toxins to cells is necessary before the toxic effects take place. A number of strains of *Escherichia coli* contain a mannose-binding lectin, which binds to mannose-rich portions of glycoproteins on the epithelial cells of the intestine. Cholera toxin is a protein with two subunits. One subunit carries the toxic components, which have no activity until they enter the cells, while the other subunit binds to the gangliosides on intestinal epithelial cells, after which the toxic subunit penetrates the cell.

The binding of influenza virus to host cells is a prerequisite for the initiation of infection. This is mediated by the influenza haemagglutinin, a lectin that is specific for the recognition of sialic acid and is located on the viral surface. Binding of this viral lectin is to the sialylated glycoconjugates on cell surfaces.

One of the first animal lectins to be identified was the hepatic lectin. Serum glycoproteins bear sialic acid residues at the nonreducing termini of the glycan chains and circulate normally in the bloodstream. Removal of this sugar from the chains exposes D-galactose at the new reducing termini. The resulting asialoglycoproteins are rapidly removed from circulation owing to their uptake in liver cells by the hepatic lectin.

Cell–cell interactions are exemplified by the homing of leucocytes to sites of inflammation. Inflammation is the body's reaction to physical damage or invasion by an infectious reagent or antigen. Leucocytes circulate normally and do not bind to cells until the onset of inflammation. Adhesion molecules (selectins), which are membrane bound, allow the leucocytes to interact with epithelial cells. This is achieved by recognition of the carbohydrate sequences on the surface glycoconjugates of the cells.

## Summary

Carbohydrates are the most abundant organic substances found in nature. They occur as simple sugars (monosaccharides) and in bound form as oligosaccharides, polysaccharides or glycoconjugates. In these higher forms, the monosaccharide units are joined together by glycosidic linkages.

Carbohydrates have numerous important functions in living material. In animals, glucose and glycogen are the main sources of energy. In starch, the polysaccharide is produced as a result of photosynthesis. Polysaccharides, such as cellulose, form part of the cell wall of plants.

Peptidoglycans and lipopolysaccharides form the protective outer cell wall of microorganisms.

Many mammalian tissues contain proteoglycans in which polysaccharides are attached to a protein backbone, and are components of connective tissue and skin.

Glycoconjugates are frequently located on the surface of membranes in animal cells. The carbohydrate portion of these molecules may be antigenic and, in certain cases, sequences of oligosaccharides are recognized by specific proteins (lectins) in complementary cells. This recognition mechanism is responsible for numerous cell–cell and cell–molecule interactions.

## Further Reading

- Binkley RW (2000) *Modern Carbohydrate Chemistry*. New York: Marcel Dekker.
- Bush CA, Martin-Pastor M and Imberty A (1999) Structure and conformation of complex carbohydrates of glycoproteins, glycolipids and bacterial polysaccharides. *Annual Review of Biophysics and Biomolecular Structure* **28**: 1–269.
- Kuberan B and Lindhardt RJ (2000) Carbohydrate-based vaccines. *Current Organic Chemistry* **4**: 653–677.
- Lindhorst TK (2000) *Essentials of Carbohydrate Chemistry and Biochemistry*. London: Wiley-VCH.
- Lis H and Sharon N (1998) Lectins: carbohydrate-specific proteins that mediate cellular recognition. *Chemistry Review* **98**: 637–674.